MAS DOC
Mathematics and Statistics
Centre for Doctoral Training

Point particle interactions on surface biomembranes: second order splitting and surface finite elements

by

Philip Justin Herbert

Thesis

Submitted for the degree of

Doctor of Philosophy

Mathematics Institute

The University of Warwick

September 2020

MENS AGITAT MOLEM

UNIVERSITAS WARWICENSIS

Contents

List of Tables

List of Figures

Acknowledgments

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) Grant EP/H023364/1 within the MASDOC Centre for Doctoral Training at the University of Warwick.

My many thanks go to Charlie, for introducing me to many interesting topics of mathematics, for providing me with many opportunities and also for his part in setting up MASDOC. I am grateful to Björn for always making time to chat.

I would like to thank my year of MASDOC and those I have shared an office with for making the four years of a PhD enjoyable. Particularly I would like to thank Luke, Jaro, Anna and Quirin. Many thanks to my friends outside of Warwick, in particular Josef, Andreas and Kat.

I thank the Free University of Berlin for their support to visit. Those I met in Berlin, I thank for a productive and fun time.

I am grateful to my thesis examiners, Klaus and Tobias, for their detailed feedback which has helped to improve this thesis.

I dedicate this thesis to my parents and cats to whom I am eternally grateful.

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work has been carried out by myself under the supervision of Charles M. Elliott and Björn Stinner.

Parts of this thesis have been published by the author, parts of Chapter 2 and Chapter 3 appear in a paper co-authored with Charlie Elliott, *Second order splitting of a class of fourth order PDEs with point constraints*, published in Mathematics of Computation [38]. Parts of Chapter 1 appear in an article co-authored with Charlie Elliott and Luke Hatcher, *Small deformations of spherical biomembranes*, due to appear in The Role of Metrics in the Theory of Partial Differential Equations, Volume 85 of Advanced Studies in Pure Mathematics [36]. Parts of Chapter 2, Chapter 4, Appendix A, Appendix B and Appendix C form a paper co-authored with Charlie Elliott, *A formula for membrane mediated point particle interactions on near spherical biomembranes* which has been submitted for publication [37].

Abstract

We study the well-posedness and approximation of mathematical models for small deformations of biological membranes where the deformations are due to point constraints. The differentiability of of the membrane energy with respect to the movement of the point constraints is studied. We begin by reviewing mathematical theory related to the shape of biomembranes and embedded proteins. We show that modifications of established theory hold and introduce notation which allows us to easily discuss the movement of many proteins embedded into the surface.

We then discuss the well-posedness of an abstract second order splitting method with linear constraints, which we will apply to the energy minimising biomembrane with embedded proteins. We also consider a penalty method to weakly enforce the constraints. It is shown that the solution of this penalty method converges strongly to the solution of the constrained problem. We consider the abstract numerical analysis of these problems. Numerical experiments are given, demonstrating the convergence theory presented.

After this, we consider the differentiability of the energy of the optimal membrane with point constraints with respect to a tangential movement of the points. We demonstrate that the energy is differentiable and give a convenient characterisation of the derivative which is efficient to evaluate. This numerically accessible derivative is employed in some numerical experiments.

We conclude by discussing some directions to extend the theory presented, or ideas which are highly related to the studied theory. In particular, we discuss the extension to consider small deformations of a near-tube membrane.

Chapter 1

Introduction

1.1 Biological membranes

Biomembranes are thin lipid bilayers which surround nearly all living cells. The bilayer forms a barrier between the cell and its surroundings and is typically composed of lipid molecules whose tails are hydrophobic and heads hydrophilic. Due to this hydrophobic/philic composition, a collection of lipid molecules immersed in an aqueous solution will often form a bilayer, a diagram of a portion of such a structure is depicted in Figure 1.1. The bilayer also contains many proteins.

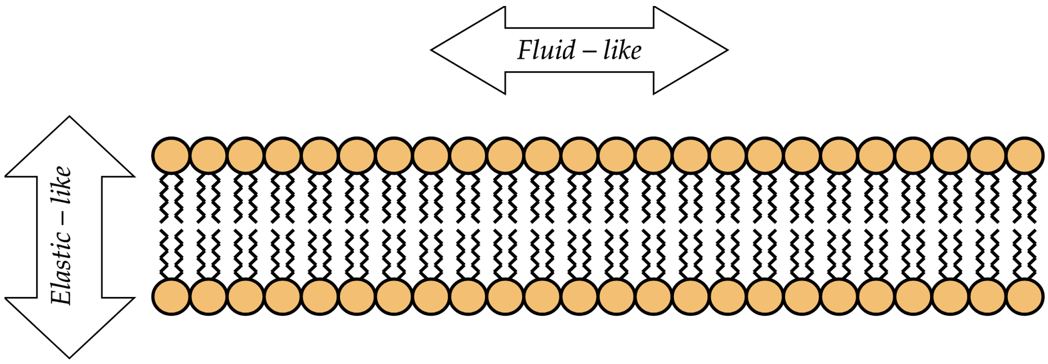
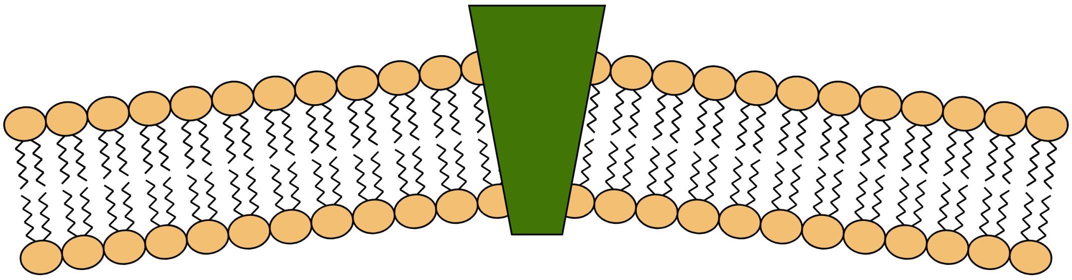
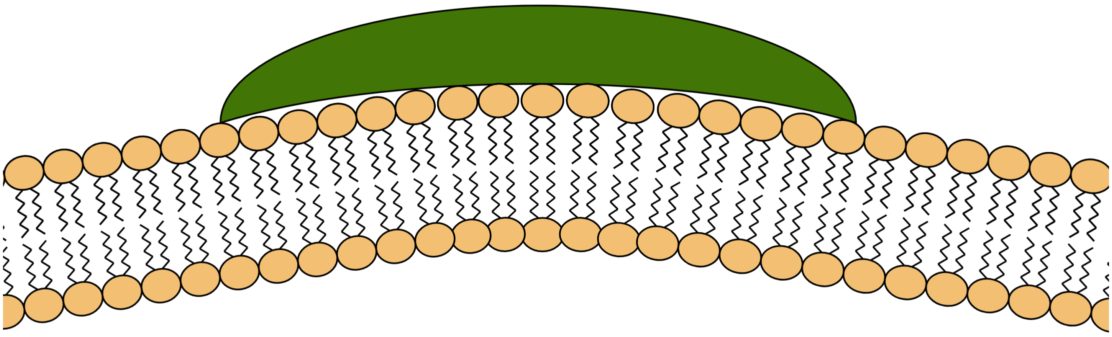


Figure 1.1: Lipid bilayer

Membrane proteins have a variety of forms and purposes. Proteins may be attached to a surface of, embedded into or span the entire bilayer. Diagrams demonstrating a protein spanning and attached to a surface of the bilayer may be found in Figure 1.2. This thesis will focus on the case of proteins attached to the membrane, rather than embedded within. It is important to note that this is not an extensive list of how a membrane may be deformed. An overview of the mechanisms of membrane deformation may be found in [67].



(a) Protein spanning the bilayer



(b) Protein attached to bilayer

Figure 1.2: Protein induced membrane deformation

The bilayer sheet has both elastic and fluid properties, these are annotated in Figure 1.1. The elastic properties relate to the energy required to bend the membrane and the fluid properties are meant in the sense that the lipids are able to rearrange themselves within the bilayer. Both of these properties are important, the elastic properties give the shape of the membrane and the fluid properties allow the proteins which are attached to the membrane to move to a desired position, for example, through membrane mediated interactions and protein-protein interactions.

1.2 Mathematical model

Established continuum models treat the biomembrane as a deformable surface of negligible thickness whose deformation is described by an energy functional depending on the curvature of the surface. The Canham-Helfrich energy [53, 15]

$$\mathcal{E}_{CH}(\mathcal{M}) := \int_{\mathcal{M}} \left(\frac{\kappa}{2}(H - c_0)^2 + \sigma + \kappa_G K \right) \mathrm{d}\mathcal{M} \tag{1.1}$$

is a classical energy, describing the equilibrium and near-equilibrium properties of biological membranes. Here $\mathcal{M}$ is a two dimensional hypersurface in $\mathbb{R}^3$ modelling the biomembrane. The quantities H and K are the mean and Gauss curvatures respectively, with $\kappa > 0$ and $\kappa_G \in \mathbb{R}$ the bending rigidities associated to the mean and Gauss curvature respectively. The surface tension is given by $\sigma \geq 0$ and finally c_0 is the bending rigidity, the preferred curvature of the membrane. We note that σ does not typically appear in the Canham-Helfrich energy, however σ may arise as a Lagrange multiplier enforcing a surface area constraint and κ_G arise in enforcing certain boundary conditions, should a boundary exist. It is worth mentioning the special case in which one takes $\kappa = 1$, $\sigma = \kappa_g = c_0 = 0$, this is know as the Willmore energy [80],

$$\mathcal{E}_W(\mathcal{M}) := \frac{1}{2}\int_{\mathcal{M}} H^2 \mathrm{d}\mathcal{M}. \tag{1.2}$$

1.2.1 Minimisation

It is common to minimise the Canham-Helfrich energy with constraints. We mention two of the most common constraints considered. For a closed membrane ($\partial\mathcal{M} = \emptyset$), one may consider the membrane to have fixed enclosed volume, which corresponds to the semi-permeability of the membrane and assuming an isotonic environment. A fixed area constraint is frequently considered, corresponding to the incompressibility of the lipid bilayer.

In recent years, the minimisation of this energy, with volume and area constraints, has attracted a lot of attention from the geometric analysis community. The article [75] finds minimisers of the Willmore energy for surfaces with Dirichlet boundary and [29] for the Canham-Helfrich energy with Dirichlet boundary. In [30], the author demonstrates lower-semicontinuity of the Canahm-Helfich energy for closed oriented varifolds. The article [68] shows existence and regularity for minimisers with the topology of a sphere. A variation of this problem, considering a multi-phase membrane energy problem has appeared in [12]; the gradient flow of this multi-phase energy has also been studied using phase fields [40, 39, 41].

1.2.2 Approximation of minimisers

A common method to approximate minimisers of energies is to consider a gradient flow of the energy. This leads to geometric evolution equations for geometric energies. Many of the studies of the flows of higher order surface energy are restricted to the case of the Willmore flow. In the simple case of the Willmore energy, it is known that, for initial data sufficiently close to the sphere, solutions converge to the sphere itself [62].

Numerical approximation of these geometric minimisation problems has flourished in recent years. A number of studies assume axi-symmetry of a surface in order to reduce the complexity, for example [6, 7, 5, 45]. A graph based method in a flat domain appears in [23], which considers a C^1 finite element method for the Willmore flow of 2-dimensional graphs. The recent article [61] gives an evolving surface finite element semi-discretisation of Willmore flow utilising mixed finite elements to approximate the normal to the surface and the mean curvature of the surface. This discretisation is shown to be stable and under appropriate assumptions on the underlying continuous problem, shown to converge.

There also have been advances in the approximation of minimisers of this energy by the use of phase fields. In the same way that the Allen–Cahn equation is known to approximate mean curvature flow, see [22, Section 7] and the references therein, many articles have used a phase field methodology to approximate Willmore flow [73, 43] with the reference [13] giving an overview of different choices of phase field energies approximating the Willmore energy.

1.2.3 Simplifications

Often, we are interested in surfaces $\mathcal{M}$ which have the same topology as the sphere. The following Gauss-Bonnet Theorem is well-known and useful for simplifying our energy.

Theorem 1.2.1. *Let $\mathcal{M}$ be a sufficiently smooth compact two-dimensional surface, with sufficiently smooth boundary $\partial\mathcal{M}$. Let K be the Gauss curvature of $\mathcal{M}$ and κ_g be the geodesic curvature of $\partial\mathcal{M}$, then*

$$\int_{\mathcal{M}} K + \int_{\partial\mathcal{M}} \kappa_g = 2\pi\chi(\mathcal{M}),$$

where $\chi(\mathcal{M})$ is the Euler characteristic of $\mathcal{M}$.

The result may be found in [26], for example. In light of the Gauss-Bonnet Theorem, when κ_G is constant and our surface has no boundary, it is permissible to neglect the Gauss curvature term. We also choose to take $c_0 = 0$. The case where $c_0 \neq 0$ is considered in [36, 35], where (1.1) is coupled to a phase field and c_0 is a function of the phase. These works derive a small deformations energy for the coupled energy and analyse its gradient flow.

1.2.4 Immersions

A typical approach is to introduce a reference surface Γ. Having a reference surface allows us to write the minimisation problem to be finding an injective function $X\colon \Gamma \to \mathcal{M} \subset \mathbb{R}^3$, rather than finding a surface. The topology of $\mathcal{M}$ is preserved by the assumption that the map X will be continuous. We write $\mathcal{M}$ to be

$$\mathcal{M} = \{X(p) : p \in \Gamma\}.$$

We will be interested in the case that

$$X = \mathrm{id}_\Gamma + \rho u \nu \tag{1.3}$$

for $\rho \ll 1$ and some function $u\colon \Gamma \to \mathbb{R}$, where ν is the unit normal to Γ. We will refer to this as the *small deformations* methodology. With these simplifications in mind, and an abuse of notation

$$\mathcal{E}(X) := \int_\Gamma \left(\frac{\kappa}{2}H^2 + \sigma\right) \mathrm{d}\mu, \tag{1.4}$$

where $\mathrm{d}\mu$ denotes the surface area element of $\mathcal{M}$ associated with the map X.

1.3 Point constraint models

As previously discussed, we are interested in the case where proteins are attached to the membrane. We say that this corresponds to requiring that the charged ends of the protein $\{z_i\}_{i=1}^L \subset \mathbb{R}^3$ are contained in the image of Γ under X, this could however be relaxed to require that $\{z_i\}_{i=1}^L$ are 'close' to the image under X. This corresponds to a *hard* constraint problem and *soft* constraint problem respectively.

We now write down two problems which motivate our studies in the case of a closed surface. Let $Z = \{z_j\}_{j=1}^L$ be a collection of disjoint points in $\mathbb{R}^3$, $\mathcal{V}(X(\Gamma))$ the volume of the domain enclosed by the image of Γ under X, $V_0 > 0$ and $\delta \in \mathbb{R}^L$ with $\delta_j > 0$ for $j = 1, ..., L$.

Problem 1.3.1. *Find $X^*_\delta \in K(\Gamma) := \{X \in C^2(\Gamma) : X$ is a diffeomorphism, $\mathcal{V}(X(\Gamma)) = V_0\}$ such that X^*_δ minimises*

$$\mathcal{E}_{\delta,Z}(X) := \mathcal{E}(X) + \sum_{j=1}^L \frac{1}{2\delta_j} d\left(X(\Gamma), z_j\right)^2,$$

where $d(X(\Gamma), y) := \inf_{x\in\Gamma} |y - X(x)|$.

The values $\delta_j > 0$ might represent the reciprocal of an elastic spring constant associated to the particle attachment. If $\delta_j = 0$, we obtain point Dirichlet constraints.

Problem 1.3.2. *Find $X^* \in K(\Gamma, Z) := \{X \in K(\Gamma) : Z \in \mathrm{Im}(X)\}$ such that*

$$\mathcal{E}(X^*) = \inf_{X\in K(\Gamma,Z)} \mathcal{E}(X).$$

The point constraints in Problem 1.3.2 are said to be hard constraints, the penalisation which appears in Problem 1.3.1 may be called soft constraints, where δ_j is interpreted as a penalisation parameter.

Problems 1.3.1 and 1.3.2 are explored in Section 2.1 through the small deformations methodology around a sphere.

1.3.1 Minimising particle configurations

Once these problems are solved for a given Z, one may consider a map $Z \mapsto X_Z$, where X_Z is a solution to Problem 1.3.2. It is then of interest to pose the following problem:

Problem 1.3.3. *Given a reference particle configuration* $\mathbf{Z} = \{Z_i\}_{i=1}^M$*, where* $Z_i = \{z_j^i\}_{j=1}^{L_i} \subset \mathbb{R}^3$*. Let* $Q \colon \mathbb{R}^6 \times \mathbb{R}^3 \to \mathbb{R}^3$ *be given by*

$$Q(q; x) = (q_1, q_2, q_3)^T + R_1(q_4)R_2(q_5)R_3(q_6)x,$$

where $R_i(\alpha)$ *is the rotation map about the* x_i *axis. Finally, for* $p = (q_1, ..., q_M) \in \mathbb{R}^{6\times M}$*, write* $\mathbf{Z}(p) := \{Q(q_i; Z_i)\}_{i=1}^M$*, where* $Q(q_i; Z_i) := \{Q(q_i; z_j^i)\}_{j=1}^{L_i}$*. Find* $p^* \in \mathbb{R}^{6\times M}$ *such that* p^* *minimises the map*

$$p \mapsto \mathcal{E}(X_{\mathbf{Z}(p)}).$$

This problem is the minimisation problem to find an energy minimising membrane-particle configuration. We note that this problem is only for minimising the membrane mediated energy, it would be natural to include pairwise particle interactions.

The above problem is given in a parameterised formulation, this parameterisation does not allow for non-rigid movements of the groups $\{Z_i\}_{i=1}^M$. It is certainly of interest to consider a relaxation where the groups are not constrained to move in a rigid manner, reflecting the fact that the proteins the groups represent need not be rigid bodies. We do not consider this relaxation, but we suggest that replacing the map X_Z with $X_{\delta,Z}$, the minimiser of Problem 1.3.1, and the energy $\mathcal{E}$ with $\mathcal{E}_{\delta,Z}$ in Problem 1.3.3 could be considered to model an elastic-like protein.

Problem 1.3.3 is explored in Chapter 4, where the gradient of the small deformations energy with respect to 'tangential changes' in particle location is calculated. The calculation of the gradient allows a gradient flow or gradient descent, which may be considered in the future.

1.4 Related literature

1.4.1 Monge-Gauge

In the work [34], the Canham-Helfrich energy (1.1) is approximated by the Monge-Gauge,

$$\mathcal{E}_{MG}(u) := \frac{1}{2}\int_\Omega \kappa\,(\Delta u)^2 + \sigma|\nabla u|^2, \tag{1.5}$$

where $\Omega \subset \mathbb{R}^2$ is an open bounded domain, with sufficiently smooth boundary and $u \colon \Omega \to \mathbb{R}$ is the height function of the membrane. The problem of minimising $\mathcal{E}_{MG}(v)$ subject to the constraints

$$v|_\gamma = h, \quad \partial_\nu v|_\gamma = s$$

is given, where γ is a closed, sufficiently smooth curve, $\partial_\nu = \nu \cdot \nabla$ is the normal derivative on γ and $h \in H^{3/2}(\gamma)$, $s \in H^{1/2}(\gamma)$ given functions. In addition, the problem where the above constraints are penalised rather than enforced is considered. Well-posedness over spaces where $\mathcal{E}_{MG}$ is coercive are presented. The article [51] considers the numerical approximation of these minimisation problems with the C^1-conforming Bogner-Fox-Schmidt finite elements.

The work [34] gives the derivative of the constrained minimum of the membrane energy $\mathcal{E}_{MG}$ with respect to the curves moving by use of shape calculus and in [50] is developed with the use of a pull back method.

1.4.2 Spherical Monge-Gauge

In the article [32], the authors derive a Monge-Gauge-like quadratic energy for the Lagrangian approximating the Canham-Helfrich energy with fixed enclosed volume for a near spherical membrane. This is done by considering the Taylor expansion of the Lagrangian in orders of ρ, for ρ as in (1.3). The energy is given as, for $\Gamma = \mathbb{S}^2(R)$, the sphere with radius $R > 0$,

$$\frac{1}{2}\int_\Gamma \kappa\left(\Delta_\Gamma u\right)^2 + \left(\sigma - \frac{2\kappa}{R^2}\right)|\nabla_\Gamma u|^2 - \frac{2\sigma}{R^2}u^2, \tag{1.6}$$

and must be accompanied by the constraint that $\int_\Gamma u = 0$. It is clear that this energy need not be positive definite, however on the subspace considered, mean-value free functions, it is shown to be positive semi-definite. On the subspace with the additional constraint $\int_\Gamma u\nu_i = 0$ for $i = 1, 2, 3$ it is shown that it is indeed coercive. With this energy, the minimisation problem is considered with point forcing and point value constraints. In the work, the authors also present a numerical approximation for the minimisation problem, the approximation is based on considering the problem as coupled second order equations. This leads to the work of [33], where the numerical method is generalised, with a particular application of the spherical Monge-Gauge with point penalties.

1.5 Surface calculus and finite elements

We recall some definitions and results from surface PDE and surface finite element methods which are necessary throughout this thesis. For full details, the reader is referred to [28].

1.5.1 Surface calculus

Let Γ be a closed C^k-hypersurface in $\mathbb{R}^3$, i.e. a surface, where k is as large as needed but at most 4 and at least 2. There is a bounded domain $U \subset \mathbb{R}^3$ such that $\partial U = \Gamma$. The unit normal ν to Γ that points out of this domain U is called the outwards unit normal. We write $P_\Gamma := I - \nu \otimes \nu$ on Γ to be, at each point on Γ, the projection onto the tangent space of Γ at that particular point, where we are writing I to be the 3×3 identity matrix. For a differentiable function $f : \Gamma \to \mathbb{R}$ we define the surface gradient by

$$\nabla_\Gamma f := P_\Gamma \nabla \overline{f},$$

where $\overline{f}$ is a differentiable extension of f to an open neighbourhood of Γ. Here ∇ denotes the standard gradient in $\mathbb{R}^3$. This given definition of the surface gradient depends only on the values of f on Γ, this is shown in [28, Lemma 2.4]. We write $\underline{D}_i f := \left(\nabla_\Gamma f\right)_i$ to be the i-th component of the surface derivative, for $i = 1, 2, 3$. The map $\mathcal{H} := \nabla_\Gamma \nu$ is called the extended Weingarten map and is symmetric with zero eigenvalue in the normal direction. The eigenvalues κ_1, κ_2 associated to the tangential eigenvectors of $\mathcal{H}$ are the principal curvatures of Γ. The product $\kappa_1\kappa_2 =: K$ gives the Gauss curvature and the sum $\kappa_1 + \kappa_2 =: H$ gives the mean curvature. The following Stokes-like theorem may be found in [28, Theorem 2.10] and has application for integration by parts formulae.

Theorem 1.5.1. *Let Γ have smooth boundary and $f \in C^1(\Gamma)$, then*

$$\int_\Gamma \nabla_\Gamma f = \int_\Gamma f H \nu + \int_{\partial\Gamma} f \mu, \tag{1.7}$$

where μ is the unit conormal vector to $\partial\Gamma$.

For a twice differentiable function, the Laplace-Beltrami operator is defined by

$$\Delta_\Gamma f := \mathrm{div}_\Gamma \nabla_\Gamma f = \sum_{i=1}^{3} \underline{D}_i \underline{D}_i f.$$

We write $D^2_\Gamma f$ to be the surface Hessian and Lemma 2.6 in [28] shows that the surface Hessian is, in general, not symmetric with the relation

$$\underline{D}_i \underline{D}_j f - \underline{D}_j \underline{D}_i f = (\mathcal{H}\nabla_\Gamma f)_j \nu_i - (\mathcal{H}\nabla_\Gamma f)_i \nu_j. \tag{1.8}$$

It is well-known [28, Lemma 2.8] that for Γ a C^2 surface, there is a small neighbourhood, $\mathcal{N}_\delta$, around Γ of width $\delta > 0$ and maps $d \colon \mathcal{N}_\delta \to \mathbb{R}$, $\pi \colon \mathcal{N}_\delta \to \Gamma$ such that for any $\tilde{X} \in \mathcal{N}_\delta$ the decomposition

$$\tilde{X} = \pi(\tilde{X}) + d(\tilde{X})\nu(\pi(\tilde{X})) \tag{1.9}$$

is unique. We call d the oriented distance function and π the closest point projection.

1.5.1.1 Surface Sobolev spaces

The Lebesgue spaces $L^p(\Gamma)$, $p \in [1, \infty)$ are defined on Γ by

$$L^p(\Gamma) := \left\{ f \colon \Gamma \to \mathbb{R} : f \text{ measurable with respect to the Hausdorff measure on } \Gamma, \int_\Gamma |f|^p < \infty \right\}.$$

For $p = \infty$ this has the usual extension where we replace the integral with the essential supremum. These spaces are equipped with the standard norms,

$$\|f\|_{L^p(\Gamma)} := \left(\int_\Gamma |f|^p \right)^{1/p},$$

with $\|f\|_{L^\infty(\Gamma)} := \operatorname{ess\,sup}_{x\in\Gamma} |f(x)|$.

We say $f \in L^1(\Gamma)$ has weak derivative $v_i = \underline{D}_i f \in L^1(\Gamma)$ if, for every $\phi \in C^1(\Gamma)$ with compact support,

$$\int_\Gamma f \underline{D}_i \phi = -\int_\Gamma \phi v_i + \int_\Gamma f \phi H \nu_i.$$

For $k \in \mathbb{N}$, $k \geq 1$, $p \in [1, \infty)$, we define

$$W^{k,p}(\Gamma) = \left\{ f \in W^{k-1,p}(\Gamma) : \underline{D}_i D_\Gamma^{k-1} f \in L^p(\Gamma),\, i = 1,2,3 \right\},$$

where $W^{0,p}(\Gamma) := L^p(\Gamma)$, D_Γ^{k-1} denotes all surface derivatives of order $(k-1)$ of f and these spaces are equipped with the expected norms:-

$$\|u\|_{W^{k,p}(\Gamma)} := \sum_{i=0}^{k} \|D_\Gamma^i u\|_{0,p}.$$

For $p = 2$ we write $H^k(\Gamma) := W^{k,2}(\Gamma)$. For $k \in \mathbb{N}_0$, $\alpha \in (0,1]$, we define the surface Hölder spaces

$$C^{k,\alpha}(\Gamma) := \left\{ u \in C^k(\Gamma) : \sup_{x,y \in \Gamma} \frac{|D_\Gamma^k u(x) - D_\Gamma^k u(y)|}{|x-y|^\alpha} < \infty \right\}.$$

We have the following embedding theorems for Γ sufficiently smooth.

Theorem 1.5.2 (Surface Sobolev embeddings)**.** *Let $k, l \in \mathbb{N}_0$, with $k > l$, let $p \in [1, \infty)$, then for $q \geq 1$ satisfying*

$$\frac{1}{p} - \frac{k}{2} = \frac{1}{q} - \frac{l}{2},$$

it holds that the embedding

$$W^{k,p}(\Gamma) \subset W^{l,q}(\Gamma)$$

is continuous, furthermore, there is $C > 0$ such that for any $u \in W^{k,p}(\Gamma)$

$$\|u\|_{W^{l,q}(\Gamma)} \leq C \|u\|_{W^{k,p}(\Gamma)}.$$

If $kp > 2$, we have that, for $r \in \mathbb{N}_0$, $\alpha \in (0,1]$ satisfying

$$\frac{1}{p} - \frac{k}{2} = -\frac{r+\alpha}{2},$$

the embedding

$$W^{k,p}(\Gamma) \subset C^{r,\alpha}(\Gamma)$$

is continuous, furthermore, there is $C > 0$ such that for any $u \in W^{k,p}(\Gamma)$

$$\|u\|_{C^{r,\alpha}(\Gamma)} \leq C \|u\|_{W^{k,p}(\Gamma)}.$$

Proof. The proof follows by locally applying the Sobolev embedding theorems [1] on the flat domain with charts. $\square$

The standard compactness results for these embeddings also hold.

Theorem 1.5.3. *Let $k, l \in \mathbb{N}$, $k > l$ and $p, q \in [1, \infty)$ with*

$$\frac{1}{p} - \frac{k}{2} < \frac{1}{q} - \frac{l}{2}.$$

Then the embedding

$$W^{k,p}(\Gamma) \subset W^{l,q}(\Gamma)$$

is compact.

1.5.2 Surface finite elements

We assume that the surface Γ is approximated by a polyhedral hypersurface

$$\Gamma_h = \bigcup_{K \in \mathcal{T}_h} K,$$

where $\mathcal{T}_h$ is a set of two-dimensional simplices in $\mathbb{R}^3$ which form an admissible triangulation. For $K \in \mathcal{T}_h$ the diameter of K is $h(K)$ and the radius of the largest (2-dimensional) ball contained in T is $\rho(K)$. Set $h := \max_{K \in \mathcal{T}_h} h(K)$ and assume the ratio between h and $\rho(K)$ is bounded independently of h. We assume that Γ_h is contained within the narrow strip $\mathcal{N}_\delta$. We assume that the restriction $\pi|_{\Gamma_h}$ of π to the polyhedral surface is a bijection between Γ_h and Γ. In addition the vertices of $K \in \mathcal{T}_h$ should lie on Γ.

The piecewise affine Lagrange finite element space on Γ_h is

$$\mathcal{S}_h := \{\chi \in C(\Gamma_h) \,:\, \chi|_K \in P^1(K) \, \forall K \in \mathcal{T}_h\},$$

where $P^1(K)$ is the set of polynomials of degree 1 or less on K. The Lagrange basis functions ϕ_i of this space are uniquely determined by their values at the so-called Lagrange nodes q_j, that is $\phi_i(q_j) = \delta_{ij}$. The associated Lagrange interpolation for a continuous function f on Γ_h is defined by

$$I_h f := \sum_i f(q_i)\phi_i.$$

We now introduce the lifted discrete spaces. We use the standard lift operator as constructed in [28, Section 4.1]. The lift f^l of a continuous function $f \colon \Gamma_h \to \mathbb{R}$ onto Γ is defined by

$$f^l(x) := (f \circ \pi|_{\Gamma_h}^{-1})(x)$$

for all $x \in \Gamma$. The inverse map g^{-l} for a continuous function $g \colon \Gamma \to \mathbb{R}$ onto Γ_h is given by

$g^{-l} := g \circ \pi$. The lifted finite element space is

$$\mathcal{S}_h^l := \{\chi^l \,|\, \chi \in S_h\}.$$

With the lifted Lagrange interpolation $I_h^l \colon C(\Gamma) \to \mathcal{S}_h^l$ given by $I_h^l(f) := (I_h f^{-l})^l$.

1.6 Structure of the thesis

We begin in Chapter 2 by introducing theory related to the minimisation of a spherical membrane problem with point constraints. Well-posedness over a larger space is given and regularity of the solution is also shown. We establish notation which allows for groups point of constraints, which will be particularly useful in Chapter 4. In Chapter 3 we consider the well-posedness and approximation of an abstract saddle point problem, which we apply to the minimisation of the spherical Monge-Gauge (1.6) with point constraints and point penalties. The case of the classical Monge-Gauge (1.5) is also considered. Numerical examples are presented, verifying the approximation results. In Chapter 4, we construct a formula for the derivative of the minimum of the membrane energy with point constraints, where the derivative is taken with respect to a parameterised position of the point constraints on the reference sphere. Numerical experiments are presented, making use of the numerical methods of the previous chapter. The experiments give a brief experimental convergence analysis of the functional and also compare the formula we construct to appropriate difference quotients. We conclude in Chapter 5 by suggesting directions in which the work presented in this thesis could be extended, with a particular emphasis on extending the derivation of [32] to include a membrane tube.

Chapter 2

A spherical membrane problem with point constraints

We start by showing well-posedness and regularity for a spherical membrane problem. We begin with the deformation model for the membrane along with a model for the particles and their attachment to the membrane.

2.1 A near-spherical biomembrane with point constraints

We now set $\Gamma = \mathbb{S}^2(R)$, the sphere of radius R centred at the origin, for some fixed $R > 0$. The model is based on the deformations of the membrane due to small external forcing. A full derivation may be found in [32], see also [36].

For the moment, we fix $K \in \mathbb{N}$ and distinct $X_i \in \Gamma$ for $i = 1, ..., K$. We set $\kappa > 0$ and $\sigma \geq 0$. First we define the energies and then give the hard constraint (Lagrange) problem and a soft constraint (penalty) problem.

Definition 2.1.1. *We define* $a \colon H^2(\Gamma) \times H^2(\Gamma) \to \mathbb{R}$ *by*

$$a(u, v) := \int_\Gamma \kappa \Delta_\Gamma u \Delta_\Gamma v + \left(\sigma - \frac{2\kappa}{R^2}\right) \nabla_\Gamma u \cdot \nabla_\Gamma v - \frac{2\sigma}{R^2} uv,$$

$T \colon C(\Gamma) \to \mathbb{R}^K$ *by*

$$Tu = (u(X_i))_{i=1}^K,$$

and for any $\epsilon > 0$, $a_\epsilon \colon H^2(\Gamma) \times H^2(\Gamma) \to \mathbb{R}$ *by*

$$a_\epsilon(u, v) := a(u, v) + \frac{1}{\epsilon}(Tu, Tv)_{\mathbb{R}^K}.$$

In addition, we set

$$J(u) := \frac{1}{2}a(u,u). \tag{2.1}$$

We notice that T is well defined by Sobolev embeddings, Theorem 1.5.2. Over $H^2(\Gamma)$, neither a_ϵ nor a are necessarily coercive, however, in [32] it is seen that they are coercive over $\{1, \nu_1, \nu_2, \nu_3\}^\perp$, where $\nu = \frac{x}{R}$ is the unit normal to Γ and $\perp$ is meant in the sense of $H^2(\Gamma)$. This is a consequence of the fact that $1, \nu_1, \nu_2, \nu_3$ are eigenfunctions of $-\Delta_\Gamma$. Furthermore, under suitable conditions on the location of the points $\{X_i\}_{i=1}^K$ we show in the following proposition that both a and a_ϵ are coercive over $\{1\}^\perp$. We use the following notation, $U := \{v \in H^2(\Gamma) : \int_\Gamma v = 0\}$ and $U_0 := \{v \in U : Tv = 0\}$.

Proposition 2.1.2. *Suppose $K \geq 4$ and $\{X_i\}_{i=1}^K$ do not lie in the same plane. Then there is $\epsilon_0 > 0$ and $C > 0$ such that for any $\epsilon < \epsilon_0$*

$$\begin{aligned} a_\epsilon(\eta,\eta) \geq& C\|\eta\|_{2,2}^2 \quad \forall \eta \in U, \\ a(\eta,\eta) \geq& C\|\eta\|_{2,2}^2 \quad \forall \eta \in U_0. \end{aligned}$$

Proof. We notice that for $u \in U_0$, it holds that $a_\epsilon(u,u) = a(u,u)$, thus we need only show the first result. In [56, Proposition 4.4.2], it is shown that a_ϵ is coercive over $\big(\mathrm{Sp}\{1, \nu_1, \nu_2, \nu_3\} \cap \mathrm{Ker}(T)\big)^\perp \cap U$. Thus it is sufficient to show that $(\mathrm{Sp}\{1, \nu_1, \nu_2, \nu_3\} \cap \mathrm{Ker}(T)) = \{0\}$. Let $v \in \mathrm{Sp}\{1, \nu_1, \nu_2, \nu_3\} \cap \mathrm{Ker}(T)$, one has that

$$v = \alpha_0 + \sum_{j=1}^{3} \alpha_j \nu_j.$$

By making note that $\nu_j(x) = \frac{x_j}{R}$ we see that v is an affine function. The condition $v \in \mathrm{Ker}(T)$ gives that the X_i are in the zero level set of the affine function v. Thus the points must lie in the same plane or $v \equiv 0$. $\square$

For the moment, we assume that $\mathcal{C} = \{X_i\}_{i=1}^K$ do not lie in the same plane. Notice that for $f = 0$, the following problems are the membrane problems in [32, 36].

Problem 2.1.3. *Given $f \in (H^2(\Gamma))^*$, find $u \in U$ minimising $J(u) - \langle f, u\rangle$ subject to $u(X_i) = Z_i$ for $i = 1, ..., K$. This has the variational formulation of finding $u \in U$ such that $u(X_i) = Z_i$ for $i = 1, ..., K$ and*

$$a(u,\eta) = \langle f, \eta\rangle \quad \forall \eta \in U_0.$$

Problem 2.1.4. *Given $f \in (H^2(\Gamma))^*$, find $u^\epsilon \in U$ minimising $J(u^\epsilon) + \frac{1}{2\epsilon}|Tu^\epsilon - Z|^2 - \langle f, u^\epsilon\rangle$. This has the variational formulation of finding $u^\epsilon \in U$ such that*

$$a_\epsilon(u^\epsilon, \eta) = \frac{1}{\epsilon}(Z, T\eta)_{\mathbb{R}^K} + \langle f, \eta\rangle \quad \forall \eta \in U.$$

Theorem 2.1.5. *There are unique solutions to both Problems 2.1.3 and 2.1.4.*

Proof. This is an application of Lax-Milgram with the coercivity of the bilinear forms shown in Proposition 2.1.2. □

In order to write down the PDE associated to these problems we need to extend the variational formulation to be posed over the whole of $H^2(\Gamma)$. Standard arguments yield that the solution of Problem 2.1.3 solves

$$a(u,\eta) + (\lambda, T\eta)_{\mathbb{R}^K} + \overline{p}\int_\Gamma \eta = \langle f, \eta\rangle \quad \forall \eta \in H^2(\Gamma).$$

and

$$a(u,\eta) + (\lambda, T\eta)_{\mathbb{R}^K} = \langle f, \eta\rangle \quad \forall \eta \in U.$$

Thus by considering smooth test functions, this gives the distributional PDE

$$\begin{aligned}\kappa\Delta_\Gamma^2 u - \left(\sigma - \frac{2\kappa}{R^2}\right)\Delta_\Gamma u - \frac{2\sigma}{R^2}u + \overline{p} + \sum_{i=1}^K \lambda_i \delta_{X_i} =& f \text{ in } \Gamma,\\ \int_\Gamma u = 0,\ u(X_i) =& Z_i \text{ for } i = 1, ..., K.\end{aligned}$$

Remark 2.1.6. *For the variational problem with penalty, the variational formulation over the of $H^2(\Gamma)$ follows similarly, yielding the distributional PDE,*

$$\begin{aligned}\kappa\Delta_\Gamma^2 u^\epsilon - \left(\sigma - \frac{2\kappa}{R^2}\right)\Delta_\Gamma u^\epsilon - \frac{2\sigma}{R^2}u + \overline{p}^\epsilon + \frac{1}{\epsilon}\sum_{i=1}^K u^\epsilon\delta_{X_i} =& f + \frac{1}{\epsilon}\sum_{i=1}^K Z_i\delta_{X_i} \text{ on } \Gamma,\\ \int_\Gamma u^\epsilon =& 0.\end{aligned}$$

It is useful to note the following.

Proposition 2.1.7. *The unique solution of Problem 2.1.3, u, satisfies $u \in W^{3,p}(\Gamma)$, $p \in (1,2)$ and is given by*

$$u = u_f + \sum_{k=1}^K Z_k\phi_k.$$

where

- *For $k = 1, ..., K$, the unique $\phi_k \in U$ $\phi_k(X_j) = \delta_{jk}$ for $j = 1, ..., K$ and*

$$a(\phi_k, \eta) = 0 \quad \forall \eta \in U_0$$

  and the unique $u_f \in U_0$ such that

$$a(u_f, \eta) = \langle f, \eta\rangle \quad \forall \eta \in U_0.$$

- *For each $k = 1, ..., K$,*

$$\lambda_k = \langle f, \phi_k \rangle - a(u_f, \phi_k) - \sum_{j=1}^{K} Z_j a(\phi_j, \phi_k)$$

and

$$\overline{p} = \langle f, \phi_0 \rangle - a(u_f, \phi_0) - \sum_{k=1}^{K} Z_k a(\phi_k, \phi_0)$$

where ϕ_0 uniquely satisfies $\phi_0(X_j) = 0$ for $j = 1, ..., K$, $\int_\Gamma \phi_0 = 1$,

$$a(\phi_0, \eta) = 0 \quad \forall \eta \in U_0.$$

Proof. The formulae for the solution are easily verified. Since $\overline{p}, \lambda$ are bounded in terms of the data, regularity for this fourth order equation on the sphere yields $-\Delta_\Gamma u \in W^{1,p}(\Gamma), p \in (1,2)$, following the arguments for a flat domain [17, 56], a full proof of the regularity result may be found in Appendix C. □

Remark 2.1.8.

- *We note that u^ϵ, the solution of the penalty problem, Problem 2.1.4, also has this regularity, that is $u^\epsilon \in W^{3,p}(\Gamma)$ for $p < 2$, which follows by an almost identical argument.*
- *In Chapter 3, we will show that $u^\epsilon \to u$ as $\epsilon \to 0$ in $W^{3,p}(\Gamma)$, for any $p < 2$.*
- *When we construct the derivative in Chapter 4, the fact the solution of Problem 2.1.3 has three weak derivatives will be used to give a more convenient representation.*
- *A related problem has been considered in* [32]*, where the authors consider the minimisation over a smaller space which enforces a fixed centre of mass for the membrane.*
- *The works* [34, 50, 51] *consider a larger solution space whereby the particles may, in some sense, tilt. The problem for this tilting on a sphere, or general domain, is of interest and may be studied in future work.*
- *An example of non-uniqueness for $K > 4$ point constraints would be to consider the location of the constraints to be contained in a plane through the axis, for example $\mathcal{C} \subset \{x \in \Gamma : x_1 = 0\}$. Then for u a solution of Problem 2.1.3 with $f = 0$, we see that $u + \alpha\nu_1 \in U$ and $J(u + \alpha\nu_1) = J(u)$ for any $\alpha \in \mathbb{R}$.*

2.1.1 A single particle model

Now we have established the well-posedness for the problem with point constraints, we wish to introduce notation to make it possible to describe multiple groups of point constraints. We wish to model the attachment of proteins to a biomembrane. A protein is considered to be a rigid discrete structure which is attached to the membrane at a finite number of fixed points. An example would be a protein such as FCHo2 F-BAR domains, where it is understood that a small number of atoms are more likely to attach to the membrane [54, 55]. The protein-biomembrane interaction is modelled by attachment at these points. This is in contrast to the case mainly considered in [51, 50], where the protein is modelled as being embedded in the membrane and attached along a curved boundary.

To begin, we restrict ourselves to a single protein in order to establish notation. We describe the protein by a finite set of distinct points $\mathcal{G} := \{\tilde{X}_i \in \mathbb{R}^3,\ i = 1, ..., M\}$. The points of $\mathcal{G}$ correspond to charged ends of the protein which attach to the membrane. The attachment constraint is the requirement that $\mathcal{G}$ is contained in $\mathcal{M}(u)$, the graph of u over Γ, which we write as

$$\mathcal{G} \subset \mathcal{M}(u) \tag{2.2}$$

and $\mathcal{M}(u) := \{p + \nu(p)u(p) : p \in \Gamma\}$. It follows that any $\tilde{X} \in \mathcal{G}$ may be uniquely decomposed into

$$\tilde{X} = \pi(\tilde{X}) + d(\tilde{X})\nu(\pi(\tilde{X})) = R\frac{\tilde{X}}{|\tilde{X}|} + \left(|\tilde{X}| - R\right)\frac{\tilde{X}}{|\tilde{X}|}$$

and the condition (2.2) becomes

$$u(\pi(\tilde{X})) = d(\tilde{X}) \quad \forall \tilde{X} \in \mathcal{G}. \tag{2.3}$$

For ease of notation, we write $X := \pi(\tilde{X})$, $z := d(\tilde{X})$ and index the points of $\mathcal{G}$ so that $\{\tilde{X}_i\}_{i=1}^M = \mathcal{G}$, hence we may write (2.3) as

$$u(X_i) = z_i \quad \forall i = 1, ..., M. \tag{2.4}$$

Definition 2.1.9. *We write $\mathcal{C} := \{\pi(\tilde{X}) : \tilde{X} \in \mathcal{G}\} = \{X_i\}_{i=1}^M$ to be the sites of attachment. Furthermore, we write*

$$u|_{\mathcal{C}} = Z$$

to be shorthand for (2.4).

2.1.2 Parametrisation of a single particle

We now parameterise the movement of a single particle. We attempt to keep our notation as similar as possible to that of [50] which deals with the rigid movement of curves in a flat domain, in contrast to our points which move on a sphere.

The assumption that the protein is rigid is meant in the sense that any movement of $\mathcal{G}$ should preserve the orientation and the distance between points. There are 6 degrees of freedom by which $\mathcal{G}$ can be moved, this is translation and rotation. We further restrict to lateral (i.e. tangential) movement of $\mathcal{G}$ over the membrane. This means that the height of attachment above

Γ, the values Z, will be independent of any movement. In the flat setting these lateral movements correspond to rotation perpendicular to the plane and translation within the plane. Although this is a strong restriction to make to the full model, it is important in this setting to avoid the particle moving out of the graph-like description.

The configuration of a single particle $\mathcal{G}$ is defined by a rigid transformation from a fixed position. We associate one point $X_\mathcal{G} \in \Gamma$ with $\mathcal{G}$. We call $X_\mathcal{G}$ the *centre* of $\mathcal{G}$. The configuration of the particle is defined by a rotation about the axis defined by $\nu(X_\mathcal{G})$ together with a *tangential translation* of $X_\mathcal{G}$ along the surface of Γ. A rotation around $\nu(X_\mathcal{G})$ is characterised by an angle, $\alpha \in \mathbb{R}$. A tangential translation is characterised by a tangent vector $\tau \in T_{X_\mathcal{G}}\Gamma \cong \mathbb{R}^2$. For this tangent vector, the idea is to consider the transport of $X_\mathcal{G}$ along the geodesic defined by τ and that the other points should follow with a rigid transformation. In the setting of a sphere, this corresponds to rotating the points by angle $|\tau|$ in the axis perpendicular to both $\nu(X_\mathcal{G})$ and τ. Thus for a particle with centre $X_\mathcal{G}$ we write $\mathcal{G}(p)$, $p = (\alpha, \tau)$ to be as described above, leading to the following definition of particle configuration.

Definition 2.1.10. *Given particle $\mathcal{G} \subset \mathbb{R}^3$ with centre $X_\mathcal{G}$ and $p = (\alpha, \tau) \in \mathbb{R} \times T_{X_\mathcal{G}}\Gamma$, we write*

$$\mathcal{G}(p) := \{\phi(p, \tilde{X}) : \tilde{X} \in \mathcal{G}\},$$

with

$$\phi(p, x) := R_T(\tau) R_n(\alpha) x \quad \forall x \in \mathbb{R}^3, \tag{2.5}$$

where $R_n(\alpha)$ is given by

$$R_n(\alpha)x := (\nu(X_\mathcal{G}) \otimes \nu(X_\mathcal{G}))x + \cos(\alpha)(\nu(X_\mathcal{G}) \times x) \times \nu(X_\mathcal{G}) + \sin(\alpha)(\nu(X_\mathcal{G}) \times x),$$

and for $\tau \neq 0$, define $\tilde{\tau} := \nu(X_\mathcal{G}) \times \frac{\tau}{|\tau|}$, $R_T(\tau)$ is given by

$$R_T(\tau)x := (\tilde{\tau} \otimes \tilde{\tau})x + \cos(|\tau|)(\tilde{\tau} \times x) \times \tilde{\tau} + \sin(|\tau|)(\tilde{\tau} \times x),$$

and $R_T(0)x = x$. A diagram showing the transformations R_n and R_T may be found in Figure 2.1. Furthermore, write

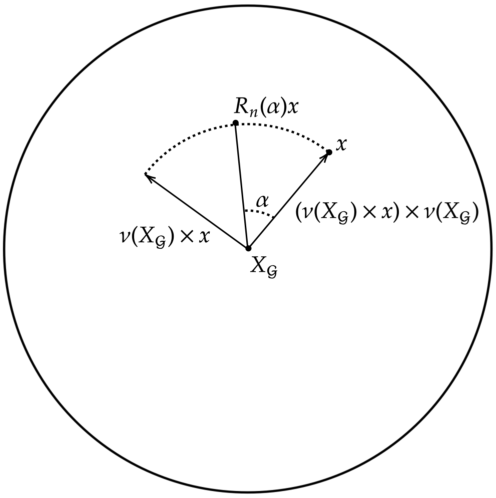
$$\mathcal{C}(p) := \{\phi(p, X) : X \in \mathcal{C}\},$$

this coincides with the projection of $\mathcal{G}(p)$ onto Γ.

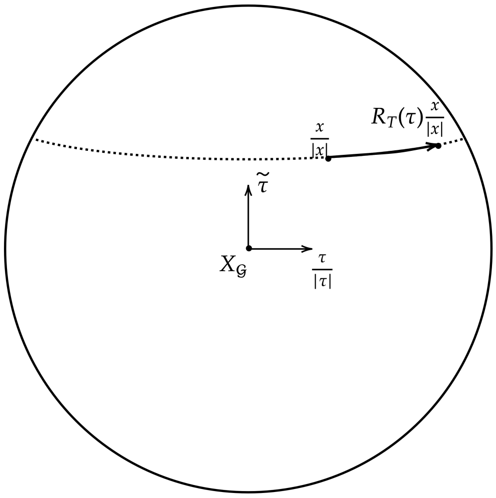
Remark 2.1.11. *The choice that $\phi(p, x) := R_T(\tau)R_n(\alpha)x$ rather than $R_n(\alpha)R_T(\tau)x$ is arbitrary. It is clear that they will both generate the same family of configurations.*

We notice $\mathcal{G} = \mathcal{G}(0)$ and similarly $\mathcal{C} = \mathcal{C}(0)$. We further note that p is periodic in the following sense. For $p = (\alpha, \tau)$, $\overline{p} = (\alpha + 2\pi, \tau)$ and $\tilde{p} = (\alpha, \tau + 2\pi\frac{\tau}{|\tau|})$ it holds,

$$\phi(p, \cdot) \equiv \phi(\overline{p}, \cdot) \equiv \phi(\tilde{p}, \cdot).$$



(a) Diagram for R_n.



(b) Diagram for R_T.

Figure 2.1: Diagrams demonstrating the transformations R_n and R_T, both with $\nu(X_\mathcal{G})$ coming out of the page.

Further note that if $\mathcal{G}$ contains only one point, $\tilde{X}_1$, and one sets $X_\mathcal{G} = X_1$, it is seen that α becomes a redundant parameter.

2.1.3 Configuration of particles

We now make the extension to multiple groups of particles.

Definition 2.1.12. *Given discrete sets with finite number of points,*

$$\mathcal{G}_1, ..., \mathcal{G}_N \subset \mathcal{N}_\delta \subset \mathbb{R}^3,$$

we write

$$\mathcal{C}_i := \{\pi(\tilde{X}) : \tilde{X} \in \mathcal{G}_i\} \text{ for } i = 1, ..., N,$$

the projection of $\mathcal{G}_i$ onto Γ. Let the $\mathcal{G}_1, ..., \mathcal{G}_N$ have centres $X_{\mathcal{G}_1}, ..., X_{\mathcal{G}_N}$ and let $p = (p_1, ..., p_N) \in \prod_{i=1}^N (\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma)$ where $p_i = (\alpha_i, \tau_i) \in \mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma$ we define

$$\phi_i(p, x) := R_{T_i}(\tau_i) R_{n_i}(\alpha_i) x \quad \forall x \in \mathbb{R}^3,$$

where the operators $R_{T_i}(\tau_i)$, $R_{n_i}(\alpha_i)$ are defined relative to the centres $X_{\mathcal{G}_i}$, as in Definition 2.1.10.

Further define

$$\mathcal{G}_i(p) := \{\phi_i(p, \tilde{X}) : \tilde{X} \in \mathcal{G}_i\} \text{ for } i = 1, ..., N,$$

and

$$\mathcal{C}_i(p) := \{\phi_i(p, X) : X \in \mathcal{C}_i\} \text{ for } i = 1, ..., N,$$

the projection of $\mathcal{G}_i(p)$ onto Γ. Observe that

$$\mathcal{G}_i(0) = \mathcal{G}_i \quad \text{and} \quad \mathcal{C}_i(0) = \mathcal{C}_i, \quad i = 1, ..., N$$

where the operators $R_{T_i}(\tau_i)$, $R_{n_i}(\alpha_i)$ are defined relative to the centres $X_{\mathcal{G}_i}$, as in Definition 2.1.10.

Definition 2.1.13. *We define the set of feasible particle configurations to be*

$$\Lambda^\circ := \left\{ p \in \prod_{i=1}^{N} (\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma) : \forall i,\, j = 1, ..., N,\ i \neq j,\ \mathcal{C}_i(p) \cap \mathcal{C}_j(p) = \emptyset \right\}.$$

We define the closure of the set of feasible particle configuration by $\Lambda := \overline{\Lambda^\circ}$. Furthermore, for $p \in \Lambda$ we define

$$\Gamma(p) := \Gamma \setminus \bigcup_{i=1}^{N} \mathcal{C}_i(p).$$

We first note that $0 \in \prod_{i=1}^{N} (\mathbb{R} \times T_{X_{\mathcal{G}_i}})$ is not a distinguished configuration. Given any non-overlapping initial configuration of particles $\{\mathcal{C}_i\}_{i=1}^N$, it is clear that Λ° is the set of all possible configurations of particles which have been moved by the rigid motions parameterised by p described at the start of Section 2.1.2.

Remark 2.1.14. *Notice that for $p \in \Lambda^\circ$, it may hold that the 'interiors' of particles overlap. As such one might want to consider a subset of Λ° whereby one defines an appropriate interior of particles and assumes that the intersection of these is empty, or perhaps one may also assign a 'radius' to each particle and consider the set where there are no points from another particle which lie inside this radius. Two ideas of these exclusion areas are shown in Figure 2.2. In this diagram, the clear dot is the centre of a particle and the black dots are the points of the particle and the exclusion area is signified by the hatched lines. The choice of this subset is not of importance when constructing the derivative, but is important when considering which particle configurations are admissible. Requiring that the particles do not overlap could be included as part of a Lennard-Jones potential, see (4.18) in [34], where it could be seen that this discussion pertains to a choice of the distance function in their formula.*

For each $p \in \Lambda^\circ$ we have a set of point constraints on elements of $H^2(\Gamma)$. This motivates the following parameterised trace operators.

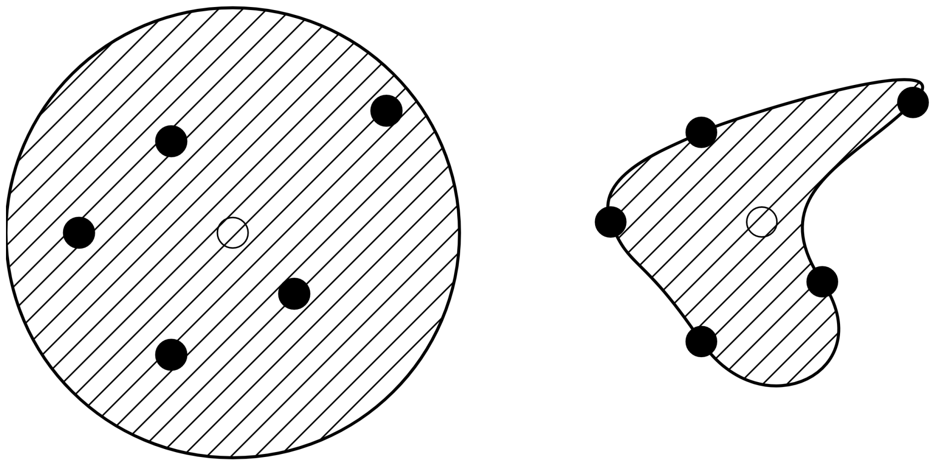


Figure 2.2: Diagram indicating the different areas which might be excluded from having part of another particle in for two idential particles. The left is the radius approach, on the right the area is given by the interior of a curve passing through all the points.

Definition 2.1.15. *Given $p \in \Lambda^\circ$:-*

- *For $i = 1, ..., N$, define the maps $T_i(p)\colon H^2(\Gamma) \to \mathbb{R}^{|\mathcal{C}_i|}$ by*

$$T_i(p)\colon v \mapsto (v \circ \phi_i(p, \cdot))\,|_{\mathcal{C}_i},$$

where $\phi_i(p, \cdot)|_{\mathcal{C}_i}$ is meant as in Definition 2.1.9.

- *For $v \in H^2(\Gamma)$, $Z \in \prod_{i=1}^N \mathbb{R}^{|\mathcal{C}_i|}$, we say $T(p)v = Z$ when*

$$T_i(p)v = Z_i \in \mathbb{R}^{|\mathcal{C}_i|} \text{ for } i = 1, ..., N$$

where Z is given by the particles $\mathcal{G}_1, ..., \mathcal{G}_N$.

- *Define the following subsets of $H^2(\Gamma)$*

$$U(p) := \{v \in U : T(p)v = Z\},$$

$$U_0(p) := \{v \in U : T(p)v = 0\}.$$

Assumption 2.1.16. *Henceforth, we assume that there is l, $1 \leq l \leq N$, such that $\mathcal{C}_l$ is not coplanar.*

Definition 2.1.17 (Membrane configurational energy)**.** *Given $p \in \Lambda^\circ$, we define $u(p) \in U(p)$ by*

$$u(p) := \operatorname*{argmin}_{v \in U(p)} J(v)$$

and we define the membrane configurational energy $\mathcal{E} : \Lambda^\circ \to \mathbb{R}$ *by*

$$\mathcal{E}(p) := J(u(p)).$$

It is clear that, by a trivial extension to Theorem 2.1.5 and Proposition 2.1.7, $u(p)$ exists, is unique and satisfies $u(p) \in W^{3,2-\delta}(\Gamma)$ for any $\delta \in (0, 1)$. For $p \in \partial\Lambda^\circ$ we do not necessarily have that a $u(p)$ exists, this is due to $U(p)$ possibly being empty.

Remark 2.1.18. *Notice that* $\mathcal{E}$ *may not be the total energy associated to the particle-membrane configuration. For example,* $\mathcal{E}$ *may be augmented with a pairwise interaction between particles modelling the electrostatic forces between different particles.*

Chapter 3

Second order splitting for the membrane problem

3.1 Introduction

We study the well-posedness and approximation of a saddle point problem posed in reflexive Banach spaces with a constraint in a Hilbert space. Let X, Y be reflexive Banach spaces, $X_0 \subset X$ be a linear subspace and S a Hilbert space with $T\colon X \to S$ being a given linear map. The problem we are interested in is:-

Given $(f, g, s) \in X^* \times Y^* \times S$ find $(u, w) \in X \times Y$ such that

$$\begin{aligned} c(u, \eta) + b(\eta, w) &= \langle f, \eta \rangle \quad \forall \eta \in X_0, \\ b(u, \xi) - m(w, \xi) &= \langle g, \xi \rangle \quad \forall \xi \in Y, \\ (Tu, z)_S &= (s, z)_S \quad \forall z \in S, \end{aligned} \tag{3.1}$$

where $c(\cdot, \cdot)$, $b(\cdot, \cdot)$ and $m(\cdot, \cdot)$ are bilinear forms and precise assumptions will be given in Section 3.2.

We approximate (3.1) by penalising the condition $(Tu - s, z)_S = 0$, rather than imposing it. This results in the problem:-

Given $(f, g, s) \in X^* \times Y^* \times S$ and $\epsilon > 0$, find $(u^\epsilon, w^\epsilon) \in X \times Y$ such that

$$\begin{aligned} c(u^\epsilon, \eta) + \frac{1}{\epsilon}(Tu^\epsilon, T\eta)_S + b(\eta, w^\epsilon) &= \langle f, \eta \rangle + \frac{1}{\epsilon}(s, T\eta)_S \qquad && \forall \eta \in X, \\ b(u^\epsilon, \xi) - m(w^\epsilon, \xi) &= \langle g, \xi \rangle && \forall \xi \in Y. \end{aligned}$$

Our abstract formulation is motivated by applications of this theory to fourth order boundary value problems arising in the modelling of biomembranes posed on a flat domain, sphere or torus, with a specific example focusing on the sphere for ease of exposition. The problems are derived in [34, 32, 51, 36] as approximations of minimisers of the Helfrich energy [53] with point constraints. In this context these arise as Dirichlet constraints on the membrane deformation modelling the attachment of point particles to the membrane at fixed locations. In the work of [14], the authors consider an optimisation problem associated with bilaplace equation with point Dirichlet conditions on a flat domain, Ω.

We have in mind the following setting. Let Γ be a curved or flat C^2 two dimensional hypersurface with or without a boundary, for $\infty > q > 2 > p > 1$, set

$$X = \left\{ \eta \in W^{1,q}(\Gamma) \, : \, \int_\Gamma \eta = 0 \right\}, \; Y = \left\{ \xi \in W^{1,p}(\Gamma) \, : \, \int_\Gamma \xi = 0 \right\}, \; L = L^2(\Gamma)$$

and $S = \mathbb{R}^N$, with bilinear forms, $c\colon W^{1,q}(\Gamma) \times W^{1,q}(\Gamma) \to \mathbb{R}$, $b\colon W^{1,q}(\Gamma) \times W^{1,p}(\Gamma) \to \mathbb{R}$ and $m\colon L^2(\Gamma) \times L^2(\Gamma) \to \mathbb{R}$. In particular we have in mind an example of the form

$$c(u, \eta) = \int_\Gamma (c_0 \nabla_\Gamma u \cdot \nabla_\Gamma \eta + c_1 u \eta), \quad b(\eta, \xi) = \int_\Gamma \nabla_\Gamma \eta \cdot \nabla_\Gamma \xi + \eta \xi, \quad m(\eta, \xi) = \int_\Gamma \eta \xi$$

where c_0, c_1 are bounded but $c(\cdot, \cdot)$ is not coercive and a linear map $T\colon W^{1,q}(\Gamma) \to \mathbb{R}^N$ defined by $(T\eta)_j := \eta(X_j)$, $j = 1, ..., N$ with $X_j \in \Gamma$, $j = 1, ..., N$.

3.1.1 Background

The study of saddle point problems is well documented, [10, 42], with many applications, for example in fluid mechanics, [47], or in linear elasticity, [3]. Note that in many of the cases in which $m \neq 0$, the authors require some strong assumptions on c, at least positive semi definite, see [10, 20, 57]. The system (3.1) is an extension to that considered in [33]. The extended system is posed over an affine subspace of $X \times Y$, rather than over the whole space. If in (3.1), we were to seach for a solution in $X_0 \times Y$, the first equation were to be considered with test functions in X_0 and the third equation to be dropped, this recovers the abstract system studied in [33]. We will use the assumptions made in [33] together with an additional assumption to handle the constraint. The assumptions will be given in Section 3.2.

In [16], the authors consider the approximation of Stokes flow by penalising the incompressibility condition. In particular, they show that the penalty terms approximate the pressure. We also consider an abstract problem with penalty and show that, in our setting, the penalty terms converge to the Lagrange multiplier associated to the constraints. Further to this, we show estimates between the solution to the problem with penalised constraint and the solution to the problem with enforced constraint. An abstract finite element theory with error bounds is presented. The results of this chapter extend those of [33] where for example, in [33, Section 6 and 7] it is shown that the well posedness theory in that paper may be applied to a problem with penalised point constraints without consideration of the convergence with respect to the penalty parameter.

The motivation for the abstract setting is to handle second order splitting for a class of fourth order surface PDEs with point Dirichlet constraints arising in the modelling of biomembranes, [32]. The setting of [33] may be directly applied to the penalty approximation for fixed penalty parameter but does not handle the hard constraint case. Here we show that the abstract setting is applicable and that the results apply to a surface finite element approximation using H^1 conforming surface finite elements. We also provide numerical experiments for this point constraint problem, considering both the grid refinements and refinements in penalty.

3.1.2 Outline of chapter

In Section 3.2 we define the abstract saddle point system with constraint and with penalty, consisting of the bilinear forms c, b, m and the inner product on the space of constraints. Well-posedness for the penalty problem trivially follows from the results of [33]. Well-posedness for the constrained problem requires additional conditions, which are natural to the standard saddle point formulation. We then show that under a set of assumptions which guarantee the problems are well posed, one obtains strong convergence with error estimates depending on the penalty parameter in the natural spaces. Two explicit examples are given in Section 3.3, the examples relate to biomembrane problems with point constraints. An abstract finite element method is then discussed in Section 3.4, which is then applied in Section 3.5 to the examples given in Section 3.3.

We conclude in Section 3.6 with some experimental examples which verify the proved convergence rates both in terms of the grid size and penalty parameter.

3.2 Abstract problem

3.2.1 Setting and problem formulation

We first define the spaces and functionals used along with the required assumptions. Throughout X, Y are reflexive Banach spaces, L is a Hilbert space with $Y \subset L$ continuously embedded and S is a separable Hilbert space with inner product $(\cdot,\cdot)_S$.

Definition 3.2.1. *Define the following*

$$\begin{aligned} c&\colon X\times X\to\mathbb{R}, \text{ bounded and bilinear,}\\ b&\colon X\times Y\to\mathbb{R}, \text{ bounded and bilinear,}\\ m&\colon L\times L\to\mathbb{R}, \text{ bounded, bilinear, symmetric and coercive,}\\ T&\colon X\to S, \text{ bounded, surjective and linear.}\end{aligned}$$

Let $s\in S$, define,

$$X_s := \{x\in X : Tx = s\}.$$

It is clear that X_s is non-empty by surjectivity of T. With this general setting in mind, we formulate a Lagrange multiplier problem and associated approximating penalised problem that we wish to consider.

Problem 3.2.2. *Given $f \in X^*$, $g \in Y^*$ and $s \in S$, find $(u, w, \lambda) \in X \times Y \times S$ such that*

$$\begin{aligned} c(u,\eta) + b(\eta, w) + (T\eta, \lambda)_S &= \langle f, \eta \rangle && \forall \eta \in X, \\ b(u,\xi) - m(w,\xi) &= \langle g, \xi \rangle && \forall \xi \in Y, \\ (Tu, z)_S &= (s, z)_S && \forall z \in S. \end{aligned}$$

Problem 3.2.3. *Given $f \in X^*$, $g \in Y^*$, $s \in S$ and $\epsilon > 0$, find $(u^\epsilon, w^\epsilon) \in X \times Y$ such that*

$$\begin{aligned} c(u^\epsilon,\eta) + \frac{1}{\epsilon}(Tu^\epsilon, T\eta)_S + b(\eta, w^\epsilon) &= \langle f, \eta \rangle + \frac{1}{\epsilon}(s, T\eta)_S && \forall \eta \in X, \\ b(u^\epsilon,\xi) - m(w^\epsilon,\xi) &= \langle g, \xi \rangle && \forall \xi \in Y. \end{aligned}$$

Remark 3.2.4. *Observe that Problem 3.2.2 is equivalent to:- Given $f \in X^*$, $g \in Y^*$ and $s \in S$, find $(u, w) \in X_s \times Y$ such that*

$$\begin{aligned} c(u,\eta) + b(\eta, w) &= \langle f, \eta \rangle && \forall \eta \in X_0, \\ b(u,\xi) - m(w,\xi) &= \langle g, \xi \rangle && \forall \xi \in Y. \end{aligned}$$

We note that the assumptions we will make for the well-posedness of these two abstract problems differ. The following assumption is required for both the Lagrange multiplier problem and the problem with penalty.

Assumption 3.2.5. *There is $C > 0$ and $\epsilon_0 > 0$ such that for all $(u, w) \in X \times Y$*

$$b(u,\xi) = m(w,\xi) \quad \forall \xi \in Y \implies C\|w\|_L^2 \le c(u,u) + \frac{1}{\epsilon_0}(Tu, Tu)_S + m(w,w). \tag{3.2}$$

3.2.2 Well posedness of Lagrange multiplier problem

Assumption 3.2.6. *There is $\zeta > 0$ such that for any $(u, w) \in X_0 \times Y$,*

$$\zeta(\|u\|_X + \|w\|_Y) \le \sup_{(\eta,\xi) \in X_0 \times Y} \frac{c(u,\eta) + b(\eta,w) + b(u,\xi) - m(w,\xi)}{\|\eta\|_X + \|\xi\|_Y}. \tag{3.3}$$

Also assume that

$$\forall (\eta,\xi) \in X_0 \times Y, \quad (\forall (u,v) \in X_0 \times Y, c(u,\eta) + b(\eta,w) + b(u,\xi) - m(w,\xi) = 0) \implies (\eta,\xi) = 0.$$

The following lemma is useful for our abstract well-posedness and may be found as a reformulation of conditions (i) and (iv) in [42, Lemma A.40], while using that S is Hilbert.

Lemma 3.2.7. *The condition that $T\colon X \to S$ is surjective is equivalent to the fact that there is $\alpha > 0$ such that*

$$\alpha \|s\|_S \leq \sup_{\eta \in X} \frac{(s, T\eta)_S}{\|\eta\|_X}.$$

When combined with the surjectivity of T, Assumption 3.2.6 is seen to be a sufficient condition for well-posedness of the standard saddle point problem:-

Find $x \in X \times Y$ and $p \in S$ such that

$$\begin{aligned} a(x, y) + d(y, p) &= \langle \tilde{f}, y \rangle \quad \forall y \in X \times Y \\ d(x, z) &= (s, z)_S \quad \forall z \in S, \end{aligned}$$

where $x = (x_1, x_2)$ with $x_1 \in X$, $x_2 \in Y$, $a(x, y) := c(x_1, y_1) + b(y_1, x_2) + b(x_1, y_2) - m(x_2, y_2)$, $\langle \tilde{f}, y \rangle = \langle f, y_1 \rangle + \langle g, y_2 \rangle$ and $d(x, z) := (Tx_1, z)_S$, which is an equivalent formulation of Problem 3.2.2.

Theorem 3.2.8. *Given Assumptions 3.2.5 and 3.2.6 hold, there is a unique solution to Problem 3.2.2. Furthermore, it holds that there is $C > 0$ such that,*

$$\|u\|_X + \|w\|_Y + \|\lambda\|_S \leq C(\|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S).$$

Proof. The existence and uniqueness of Problem 3.2.2 is a simple consequence of Assumption 3.2.6 and surjectivity of T, using a standard theorem on saddle point problems, see [42, Theorem 2.34] for example. □

3.2.3 Well posedness of penalty approximation

We will require the following assumptions on b, X and Y, as in [32].

Assumption 3.2.9. *There exist $\gamma, \beta > 0$ such that*

$$\beta \|\eta\|_X \leq \sup_{\xi \in Y} \frac{b(\eta, \xi)}{\|\xi\|_Y} \ \forall \eta \in X \ \text{and} \ \gamma \|\xi\|_Y \leq \sup_{\eta \in X} \frac{b(\eta, \xi)}{\|\eta\|_X} \ \forall \xi \in Y. \tag{3.4}$$

In addition to the above, we also require there to be sufficiently well behaved approximating spaces. This allows for a Galerkin approximation. We will see that we may pick finite element spaces satisfying the conditions.

Assumption 3.2.10. *There are finite dimensional approximating spaces $X_n \subset X$ and $Y_n \subset Y$, that is, $\forall (\eta, \xi) \in X \times Y$ there are $(\eta_n, \xi_n) \in X_n \times Y_n$ with $\|\eta - \eta_n\|_X + \|\xi - \xi_n\|_Y \to 0$. We additionally assume discrete inf-sup conditions. That is there are $\tilde{\beta}, \tilde{\gamma} > 0$, independent of n, such that*

$$\tilde{\beta} \|\eta_n\|_X \leq \sup_{\xi_n \in Y_n} \frac{b(\eta_n, \xi_n)}{\|\xi_n\|_Y} \ \forall \eta_n \in X_n \ \text{and} \ \tilde{\gamma} \|\xi_n\|_Y \leq \sup_{\eta_n \in X_n} \frac{b(\eta_n, \xi_n)}{\|\eta_n\|_X} \ \forall \xi_n \in Y_n. \tag{3.5}$$

We also assume that there is an interpolation map $I_n\colon Y \to Y_n$ for each n such that

$$b(\eta_n, I_n\xi) = b(\eta_n, \xi) \ \forall (\eta_n, \xi) \in X_n \times Y,$$

$$\sup_{\xi \in Y} \frac{\|\xi - I_n\xi\|_L}{\|\xi\|_Y} \to 0 \ \textit{as } n \to \infty.$$

We now quote two results which will be useful to refer to throughout the work, they may be found in [33, Lemma 2.1 and 2.2].

Lemma 3.2.11. *Let Assumptions 3.2.5 and 3.2.10 hold. There is a linear map $G_n\colon Y^* \to X_n$ such that for any $\theta \in Y^*$*

$$b(G_n\theta, \xi_n) = \langle \theta, \xi_h \rangle \quad \forall \xi_n \in Y_n.$$

Lemma 3.2.12. *Let Assumptions 3.2.5 and 3.2.10 hold. There is $C, N > 0$ such that for all $n \geq N$ and any $v_n \in Y_n$,*

$$C\|v_n\|_L^2 \leq c(G_n(m(v_n,\cdot)), G_n(m(v_n,\cdot))) + \frac{1}{\epsilon_0}(T(G_n(m(v_n,\cdot))), T(G_n(m(v_n,\cdot))))_S + m(v_n, v_n). \tag{3.6}$$

Now we have the required assumptions, we assert the well-posedness of Problem 3.2.3.

Theorem 3.2.13. *Given Assumptions 3.2.5, 3.2.9 and 3.2.10, there is a unique solution to Problem 3.2.3. Furthermore it holds that*

$$\|u^\epsilon\|_X + \|w^\epsilon\|_Y \leq C\left(1+\epsilon^{-1}\right)\left(\|f\|_{X^*} + \left(1+\epsilon^{-1}\right)\|g\|_{Y^*} + \epsilon^{-1}\|s\|_S\right)$$

Proof. The existence and uniqueness follows from [33, Theorem 2.2], with the estimate following from carrying through the ϵ terms. □

Recall that we are interested in the case $\epsilon \to 0$, clearly in the above estimate, the bound diverges. For numerics, one might like to take ϵ to be a function of the grid size, as such, the bounds diverging in ϵ means one must be restrictive in the relationship between grid size and ϵ. To show uniform bounds, we make use of the solution to Problem 3.2.2.

3.2.4 Convergence of penalty approximation

In this subsection we assume Assumptions 3.2.5, 3.2.6, 3.2.9, and 3.2.10 to hold.

Proposition 3.2.14. *Let (u, w, λ) solve Problem 3.2.2 and (u^ϵ, w^ϵ) solve Problem 3.2.3. Then it holds that there is $C > 0$, independent of ϵ such that*

$$\|w - w^\epsilon\|_Y + \|u - u^\epsilon\|_X + \left\|\lambda - \epsilon^{-1}T(u^\epsilon - u)\right\|_S \leq C\sqrt{\epsilon}\|\lambda\|_S.$$

Proof. From (3.4), it holds,

$$\beta\|u-u^\epsilon\|_X \le \sup_{\xi\in Y}\frac{b(u-u^\epsilon,\xi)}{\|\xi\|_Y} = \sup_{\xi\in Y}\frac{m(w-w^\epsilon,\xi)}{\|\xi\|_Y} \le C\|w-w^\epsilon\|_L,$$

where we have used the second equations of the systems. Now by taking differences of the first equations of the systems,

$$c(u^\epsilon-u,\eta)+b(\eta,w^\epsilon-w)+\frac{1}{\epsilon}(T(u^\epsilon-u),T\eta)_S = (\lambda,T\eta)_S\,\forall\eta\in X.$$

By letting $\eta=u^\epsilon-u$ in the above and from (3.2), one has,

$$\begin{aligned}C\|w-w^\epsilon\|_L^2+\left(\frac{1}{\epsilon}-\frac{1}{\epsilon_0}\right)\|T(u-u^\epsilon)\|_S^2 \le& c(u^\epsilon-u,u^\epsilon-u)+m(w^\epsilon-w,w^\epsilon-w)\\ &+\frac{1}{\epsilon}(T(u-u^\epsilon),T(u-u^\epsilon))_S\\ =&(\lambda,T(u^\epsilon-u))_S\le C\|\lambda\|_S\|T(u^\epsilon-u)\|_S\\ \le& C\frac{\epsilon}{2}\|\lambda\|_S^2+\frac{1}{2\epsilon}\|T(u^\epsilon-u)\|_S^2,\end{aligned}$$

this has shown

$$\|w-w^\epsilon\|_L+\|u-u^\epsilon\|_X+\frac{1}{\sqrt{\epsilon}}\|T(u-u^\epsilon)\|_S\le C\sqrt{\epsilon}\|\lambda\|_S, \tag{3.7}$$

which shows the result for the $\|u-u^\epsilon\|_X$.

Due to Assumption 3.2.6, one has,

$$\kappa\|w-w^\epsilon\|_Y\le \sup_{(\eta,\xi)\in X_0\times Y}\frac{c(0,\eta)+b(\eta,w-w^\epsilon)+b(0,\xi)-m(w-w^\epsilon,\xi)}{\|\eta\|_X+\|\xi\|_Y}.$$

One then calculates

$$c(u^\epsilon-u,\eta)+b(\eta,w^\epsilon-w)=0\quad\forall\eta\in X_0,$$

hence

$$\kappa\|w-w^\epsilon\|_Y\le \sup_{(\eta,\xi)\in X_0\times Y}\frac{c(u^\epsilon-u,\eta)-m(w-w^\epsilon,\xi)}{\|\eta\|_X+\|\xi\|_Y}\le C(\|u^\epsilon-u\|_X+\|w-w^\epsilon\|_L)$$

this proves the result for $\|w-w^\epsilon\|_Y$ when making use of (3.7). Finally, from surjectivity of T and Lemma 3.2.7, one has

$$\alpha\left\|\lambda-\epsilon^{-1}T(u^\epsilon-u)\right\|_S\le\sup_{\eta\in X}\frac{(\lambda-\epsilon^{-1}T(u^\epsilon-u),T\eta)_S}{\|\eta\|_X},$$

where again one calculates

$$(\lambda-\epsilon^{-1}T(u^\epsilon-u),T\eta)_S=c(u^\epsilon-u,\eta)+b(\eta,w^\epsilon-w)\quad\forall\eta\in X,$$

thus

$$\alpha \left\|\lambda - \epsilon^{-1}T(u^\epsilon - u)\right\|_S \leq C\|\lambda\|_S\sqrt{\epsilon}.$$

□

Indeed, in the case that the dual problem is well-posed, it is possible to use an Aubin-Nitsche type argument to give a higher rate of convergence for $\|u - u^\epsilon\|_X$ and $\|w - w^\epsilon\|_L$.

Proposition 3.2.15. *Let (u, w, λ) solve Problem 3.2.2, let (u^ϵ, w^ϵ) solve Problem 3.2.3. Suppose that there is $(\psi, \phi, \chi) \in X \times Y \times S$ such that*

$$\begin{aligned} c(\eta, \psi) + b(\eta, \phi) + (\chi, T\eta)_S &= 0 && \forall \eta \in X, \\ b(\psi, \xi) - m(\phi, \xi) &= (w - w^\epsilon, \xi)_L && \forall \xi \in Y, \\ (T\psi, z) &= 0 && \forall z \in S, \end{aligned} \tag{3.8}$$

and it holds that there is some $C > 0$ with

$$\|\psi\|_X + \|\phi\|_Y + \|\chi\|_S \leq C\|w - w^\epsilon\|_L.$$

Then there is a $C > 0$ independent of ϵ such that

$$\|u - u^\epsilon\|_X + \|w - w^\epsilon\|_Y \leq C\epsilon.$$

Proof. In the proof of Proposition 3.2.14, it is shown that

$$\kappa\|w - w^\epsilon\|_Y \leq C(\|u - u^\epsilon\|_X + \|w - w^\epsilon\|_L),$$

therefore it is sufficient to control $\|u - u^\epsilon\|_X + \|w - w^\epsilon\|_L$. By testing the system (3.8) with $(u - u^\epsilon, w - w^\epsilon, T(u - u^\epsilon))$ and summing the first two equations,

$$\begin{aligned} \|w - w^\epsilon\|_L^2 =& c(u - u^\epsilon, \psi) + b(u - u^\epsilon, \phi) + (\chi, T(u - u^\epsilon))_S + b(\psi, w - w^\epsilon) - m(w - w^\epsilon, \phi) \\ =& (\chi, T(u - u^\epsilon))_S. \end{aligned}$$

Where one has $b(u - u^\epsilon, \phi) = m(w - w^\epsilon, \phi)$ and $c(u - u^\epsilon, \psi) = -b(\psi, w - w^\epsilon)$ with the second following from $(T\psi, \lambda)_S = \epsilon^{-1}(T(u^\epsilon - u), T\psi)_S = 0$. Hence from the estimate for $\|T(u - u^\epsilon)\|_S$ shown in Proposition 3.2.14 along with $\|u - u^\epsilon\|_X \leq C\|w - w^\epsilon\|_L$ and $\|\chi\|_S \leq C\|w - w^\epsilon\|_L$, the result is complete. □

Remark 3.2.16. *For the dual system* (3.8) *to be well-posed, one requires the dual to Assumption 3.2.6 to hold. A sufficient condition for this is that the bilinear form c is symmetric.*

We may now use Proposition 3.2.14 to give uniform bounds on u^ϵ and w^ϵ.

Corollary 3.2.17. *Given $\epsilon > 0$ sufficiently small, let $(u^\epsilon, w^\epsilon) \in X \times Y$ solve Problem 3.2.3. Then there is $C > 0$ with*

$$\|u^\epsilon\|_X + \|w^\epsilon\|_Y + \frac{1}{\epsilon}\|T(u^\epsilon - u)\|_S \leq C(\|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S).$$

3.3 PDE example problems

In this section, we demonstrate the well-posedness of a split system to the energy discussed in Section 2.1. In addition, we demonstrate the well-posedness for the flat Monge-Gauge, both in the natural fourth order setting and a second order splitting system.

3.3.1 Second order splitting for the spherical Monge-Gauge

We will again use $\Gamma = \mathbb{S}^2(R)$, the sphere of radius $R > 0$ with N point constraints. For certain boundary value problems, splitting a fourth order equation into two second order equations is a natural approach, c.f. [31]. Here it is convenient to use an auxiliary variable $w = -\Delta_\Gamma u + u$ leading to the following coupled system holding on Γ

$$\begin{aligned} -\Delta_\Gamma w + w - \left(\frac{\sigma}{\kappa} - \frac{2}{R^2} - 2\right)\Delta_\Gamma u - \left(\frac{2\sigma}{\kappa R^2} + 1\right) u &= \frac{1}{\kappa}\left(f - \overline{p} + \Sigma_{i=1}^N \lambda_i \delta_{X_i}\right), \\ -\Delta_\Gamma u + u &= w. \end{aligned} \tag{3.9}$$

Taking $u \in W^{3,p}(\Gamma)$ as the solution of Problem 2.1.3, as in Proposition 2.1.7, we pose the first equation weakly in the dual of $W^{1,q}(\Gamma)$ with $q \in (2, \infty)$ and $u \in W^{1,q}(\Gamma)$ with $\frac{1}{p} + \frac{1}{q} = 1$. The second equation is posed in the dual of $W^{1,p}(\Gamma)$. It is clear by testing (3.9) by $\eta \in X$ and $\xi \in Y$ that the PDE system may be posed as in (3.1) using Definition 3.3.1.

Definition 3.3.1. *Let $\infty > q > 2 > p > 1$, then we define the spaces*

$$X = \left\{\eta \in W^{1,q}(\Gamma) \, : \, \int_\Gamma \eta = 0\right\},$$
$$Y = \left\{\xi \in W^{1,p}(\Gamma) \, : \, \int_\Gamma \xi = 0\right\},$$

$L = L^2(\Gamma)$ and $S = \mathbb{R}^N$, with the bilinear forms,

$$\begin{aligned} c&\colon W^{1,q}(\Gamma) \times W^{1,q}(\Gamma) \to \mathbb{R}, \\ b&\colon W^{1,q}(\Gamma) \times W^{1,p}(\Gamma) \to \mathbb{R}, \\ m&\colon L^2(\Gamma) \times L^2(\Gamma) \to \mathbb{R}, \end{aligned}$$

given by

$$
\begin{aligned}
c(u,\eta) &= \int_\Gamma \left(\frac{\sigma}{\kappa} - 2 - \frac{2}{R^2}\right) \nabla_\Gamma u \cdot \nabla_\Gamma \eta - \left(1 + \frac{2\sigma}{\kappa R^2}\right) u\eta,\\
b(\eta,\xi) &= \int_\Gamma \nabla_\Gamma \eta \cdot \nabla_\Gamma \xi + \eta\xi,\\
m(\eta,\xi) &= \int_\Gamma \eta\xi
\end{aligned}
$$

and the linear operator $T : X \to S$ by

$$T\eta := (\eta(X_1), \eta(X_2), ...\eta(X_N)),\ \ \eta \in X.$$

3.3.1.1 Verification of Assumptions 3.2.5, 3.2.9, 3.2.10 and 3.2.6

We let X_n and Y_n be the lifted discrete space $\mathcal{S}_h^l$, as defined in Section 1.5.2. More specifically, for a sequence of triangulated surfaces $(\Gamma_{h_n})_{n\in\mathbb{N}}$ with $h_n \to 0$ as $n \to \infty$ we have $X_n := X_{h_n}^l = \mathcal{S}_{h_n}^l$ and $Y_n := Y_{h_n}^l = \mathcal{S}_{h_n}^l$. We begin with the following three results which are shown in [33, Section 5].

Lemma 3.3.2. *Let $1 < p \le 2 \le q < \infty$ with p, q conjugate. There is $\beta, \gamma > 0$ such that*

$$\beta\|\eta\|_{1,q} \le \sup_{\xi\in W^{1,p}(\Gamma)} \frac{b(\eta,\xi)}{\|\xi\|_{1,p}} \quad \forall \eta \in W^{1,q}(\Gamma) \quad \text{and} \quad \gamma\|\xi\|_{1,p} \le \sup_{\eta\in W^{1,q}(\Gamma)} \frac{b(\eta,\xi)}{\|\eta\|_{1,q}} \quad \forall \xi \in W^{1,p}(\Gamma).$$

Lemma 3.3.3. *Let $1 < r \le \infty$. Then there is a bounded (independently of h) linear map $\Pi_h \colon W^{1,r}(\Gamma) \to \mathcal{S}_h^l$ given by*

$$b(\Pi_h\phi, v_h^l) = b(\phi, v_h^l) \quad \forall v_h^l \in \mathcal{S}_h^l.$$

It also holds that

$$\sup_{\psi\in W^{1,r}(\Gamma)} \frac{\|\psi - \Pi_h\psi\|_{0,2}}{\|\psi\|_{1,r}} \to 0 \quad \text{as} \quad h \to 0.$$

Lemma 3.3.4. *Let $1 < p \le 2 \le q < \infty$ with p, q conjugate. There is $\tilde\beta, \tilde\gamma > 0$ such that*

$$\tilde\beta\|\eta_h^l\|_{1,q} \le \sup_{\xi_h^l\in\mathcal{S}_h^l} \frac{b(\eta_h^l,\xi_h^l)}{\|\xi_h^l\|_{1,p}} \quad \forall \eta_h^l \in \mathcal{S}_h^l \quad \text{and} \quad \tilde\gamma\|\xi_h^l\|_{1,p} \le \sup_{\eta_h^l\in\mathcal{S}_h^l} \frac{b(\eta_h^l,\xi_h^l)}{\|\eta_h^l\|_{1,q}} \quad \forall \xi_h^l \in \mathcal{S}_h^l.$$

In order to prove well-posedness of the problem with penalty, we are left to show Assumptions 3.2.9, 3.2.10 and 3.2.5 in the appropriate spaces. We make use of Fortin's criteria with the projection to mean-value-free functions and Lemmas 3.3.2 and 3.3.4. The following Proposition gives Assumptions 3.2.9 and 3.2.10.

Proposition 3.3.5. *Assumptions 3.2.9 and 3.2.10 hold true.*

Proof. This follows from an application of Fortin's criterion [42, Lemma 4.19] to Lemmas 3.3.2 and 3.3.4 with the projection map $\overline{P}\colon u \mapsto u - \frac{1}{|\Gamma|}\int_\Gamma u$. $\square$

The final condition we need to check is the coercivity like condition, Assumption 3.2.5.

Lemma 3.3.6. *Let $\epsilon > 0$ sufficiently small and assume $(u, w) \in X \times Y$ satisfy*

$$b(u, \xi) = m(w, \xi) \quad \forall \xi \in Y.$$

Then there is $C > 0$ such that

$$C\|w\|_{0,2}^2 \leq c(u, u) + m(w, w) + \frac{1}{\kappa\epsilon}(Tu, Tu)_{\mathbb{R}^N}.$$

Proof. The condition $b(u, \xi) = m(w, \xi) \quad \forall \xi \in Y$ with $(u, w) \in X \times Y$ implies that $b(u, \xi) = m(w, \xi) \quad \forall \xi \in W^{1,p}(\Gamma)$. Elliptic regularity gives $-\Delta_\Gamma u \in W^{1,p}(\Gamma)$ and $-\Delta_\Gamma u + u = w$. Making use of this relation in $c(u, u) + m(w, w)$ gives

$$c(u, u)+m(w, w)+\frac{1}{\kappa\epsilon}(Tu, Tu)_{\mathbb{R}^N} = \int_\Gamma \left((\Delta_\Gamma u)^2 + \left(\frac{\sigma}{\kappa} - \frac{2}{R^2}\right) |\nabla_\Gamma u|^2 - \frac{2\sigma}{\kappa R^2} u^2 \right) + \frac{1}{\kappa\epsilon}(Tu, Tu)_{\mathbb{R}^N},$$

which is bounded below by $C\|u\|_{2,2}^2$, as shown in Proposition 2.1.2. Elliptic regularity applied to the condition $b(u, \xi) = m(w, \xi)$ $\forall \xi \in W^{1,p}(\Gamma)$ gives $\|w\|_{0,2} \leq C\|u\|_{2,2}$ to complete the proof. $\square$

Lemma 3.3.7. *There is $\alpha > 0$ such that for any $(u, w) \in X_0 \times Y$ it holds that*

$$\alpha(\|u\|_{1,q} + \|w\|_{1,p}) \leq \sup_{(\eta,\xi)\in X_0\times Y} \frac{c(u, \eta) + b(\eta, w) + b(u, \xi) - m(w, \xi)}{\|\eta\|_{1,q} + \|\xi\|_{1,p}}.$$

Proof. It is sufficient [42, Theorem 2.34] to show for any $(\alpha, \beta, \tilde{Z}, \tilde{f}, \tilde{g}) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}^N\times W^{1,q}(\Gamma)^* \times W^{1,p}(\Gamma)^*$, the following system has a unique solution

$$\begin{aligned} c(u, \eta) + b(\eta, w) + (\lambda, T\eta)_{\mathbb{R}^N} &= \langle \tilde{f}, \eta\rangle \quad \forall \eta \in X, \\ b(u, \xi) - m(w, \xi) &= \langle \tilde{g}, \xi\rangle \quad \forall \xi \in Y, \\ Tu &= \tilde{Z}, \\ \int_\Gamma u &= \alpha, \\ \int_\Gamma w &= \beta. \end{aligned} \tag{3.10}$$

Let $G\colon (W^{1,p}(\Gamma))^* \to W^{1,q}(\Gamma)$ be the inverse-Laplace operator such that $b(G\tilde{g}, \xi) = \langle \tilde{g}, \xi\rangle \quad \forall \xi \in W^{1,p}(\Gamma)$. We are able to find unique $u_1 \in H^2(\Gamma)$ with $Tu_1 = \tilde{Z} - TGg$, $\int_\Gamma u_1 = \alpha - \int_\Gamma Gg$ and $a(u_1, v) = \langle \tilde{f}, v\rangle - c(Gg, v)$ for all $v \in H^2(\Gamma)$ with $\int_\Gamma v = 0$ and $Tv = 0$. In particular, as discussed in Subsection 2.1,

$$a(u_1, v) + (\lambda_1, Tv)_{\mathbb{R}^n} + \overline{p}_1 \int_\Gamma v = \langle \tilde{f}, v\rangle - c(Gg, v) \quad \forall v \in H^2(\Gamma).$$

From this formulation, it then follows that $-\Delta_\Gamma u_1 \in W^{1,p}(\Gamma)$ as in Proposition 2.1.7. Defining $w_1 := -\Delta_\Gamma u_1 + u_1 + \frac{\beta - \alpha + \int_\Gamma Gg}{|\Gamma|} \in W^{1,p}(\Gamma)$, we see that we have a solution to the problem

$$\begin{aligned}
c(u_1, \eta) + b(\eta, w_1) + (\lambda_1, T\eta)_{\mathbb{R}^N} &= \langle \tilde{f}, \eta\rangle - c(Gg, \eta) \quad \forall \eta \in X,\\
b(u_1, \xi) - m(w_1, \xi) &= 0 \qquad\quad \forall \xi \in Y,\\
Tu_1 &= \tilde{Z} - TGg,\\
\int_\Gamma u_1 &= \alpha - \int_\Gamma Gg,\\
\int_\Gamma w_1 &= \beta.
\end{aligned} \tag{3.11}$$

By defining $\tilde{u} := u_1 + G\tilde{g}$ we have $(\tilde{u}, w_1, \lambda)$ is a solution to (3.10). In order to show uniqueness of the constructed solution, one may reverse the above process and appeal to the well-posedness of the fourth order equation. To do so, one uses the regularising properties of the inverse Laplacian. From the fourth order equation, one has the uniqueness of the u variable, from which the uniqueness of the other variables follows. □

3.3.1.2 Well posedness

We are now able to prove well-posedness of the following problems which represent a generalisation of Problems 2.1.3 and 2.1.4 by the inclusion of g in the right hand side of the second equation.

Problem 3.3.8. *Given $f \in (W^{1,q}(\Gamma))^*$ and $g \in (W^{1,p}(\Gamma))^*$, find $(u, w, \lambda) \in X \times Y \times \mathbb{R}^N$ such that*

$$\begin{aligned}
c(u, \eta) + b(\eta, w) + (T\eta, \lambda)_{\mathbb{R}^N} &= \langle f, \eta\rangle & \forall \eta \in X,\\
b(u, \xi) - m(w, \xi) &= \langle g, \xi\rangle & \forall \xi \in Y,\\
Tu &= Z.
\end{aligned}$$

Problem 3.3.9. *Given $f \in (W^{1,q}(\Gamma))^*$, $g \in (W^{1,p}(\Gamma))^*$ and $\epsilon > 0$, find $(u^\epsilon, w^\epsilon) \in X \times Y$ such that*

$$\begin{aligned}
c(u^\epsilon, \eta) + b(\eta, w^\epsilon) + \frac{1}{\epsilon}(Tu^\epsilon, T\eta)_{\mathbb{R}^N} &= \langle f, \eta\rangle + \frac{1}{\epsilon}(Z, T\eta)_{\mathbb{R}^N} & \forall \eta \in X,\\
b(u^\epsilon, \xi) - m(w^\epsilon, \xi) &= \langle g, \xi\rangle & \forall \xi \in Y.
\end{aligned}$$

Remark 3.3.10. *If one were to suppose that Problems 2.1.3 and 2.1.4 have right hand side* *forcing $F \in \left(H^2(\Gamma)\right)^*$, where $F = f - \Delta_\Gamma g + g$. Then, formally, the splitting $w := -\Delta_\Gamma u + u - g$ gives rise to Problems 3.3.8 and 3.3.9. This may be interpreted as a decomposition of the data F into "smooth" and "singular" components.*

Theorem 3.3.11. *There is a unique solution to Problem 3.3.9.*

Proof. The proof is an application of Theorem 3.2.13 as we have shown Assumptions 3.2.5, 3.2.9 and 3.2.10 in Propositions 2.1.2 and 3.3.5. □

We now wish to show the appropriate assumptions for the well-posedness of the following problem.

Theorem 3.3.12. *There a is unique solution to Problem 3.3.8.*

Proof. This follows directly from Lemma 3.3.7 as this proves that Assumption 3.2.6 holds true. □

From Propositions 3.2.14 and 3.2.15, we have the following result.

Corollary 3.3.13. *Let (u, w, λ) solve Problem 3.3.8 and (u^ϵ, w^ϵ) solve Problem 3.3.9. Then there is $C > 0$ such that*

$$\|u - u^\epsilon\|_X + \|w - w^\epsilon\|_L + \sqrt{\epsilon}\|w - w^\epsilon\|_Y + \sqrt{\epsilon}\|\epsilon^{-1}(Tu^\epsilon - Z) - \lambda\|_{\mathbb{R}^N} \leq C\epsilon\|\lambda\|_{\mathbb{R}^N}.$$

3.3.2 A near flat biomembrane

We give a flavour of how this same theory may be applied to the case of the Monge-Gauge. The Monge-Gauge is studied in [34] and it is noted that it is a geometric linearisation of the Canham-Helfrich energy or indeed, formally, the limit as $R \to \infty$ in the a given in Definition 2.1.1. Let Ω be a smooth bounded domain in $\mathbb{R}^2$, and fix $\kappa > 0$ and $\sigma \geq 0$. Fix $N \in \mathbb{N}$ and distinct $X_i \in \Omega$ for $i = 1, ..., N$ so that $S = \mathbb{R}^N$ and T is the evaluation map at these N points.

For this flat problem, we consider the Monge-Gauge energy [34]. The numerical analysis for this has been considered in [51] for finite size particles with constraints on closed curves using a penalty method. The authors make use of higher order H^2 conforming finite elements so do not need to split the equation.

3.3.2.1 Fourth order formulation

Definition 3.3.14. *Define $a\colon H^2(\Omega) \times H^2(\Omega) \to \mathbb{R}$ by*

$$a(u, v) := \int_\Omega \kappa \Delta u \Delta v + \sigma \nabla u \cdot \nabla v,$$

$T\colon C(\Omega) \to \mathbb{R}^N$ by

$$Tu = (u(X_i))_{i=1}^N,$$

and for any $\epsilon > 0$, $a_\epsilon \colon H^2(\Omega) \times H^2(\Omega) \to \mathbb{R}$ by

$$a_\epsilon(u,v) := a(u,v) + \frac{1}{\epsilon}(Tu, Tv)_{\mathbb{R}^N}.$$

It may be seen [34] that a is coercive over $V := H^2(\Omega) \cap H_0^1(\Omega)$, which corresponds to so called Navier boundary conditions, which we consider here.

Problem 3.3.15. *Given $f \in (H^2(\Omega))^*$, find $u \in V$ minimising $\frac{1}{2}a(u,u) - \langle f, u\rangle$ subject to $u(X_i) = Z_i$ for $i = 1, ..., N$. This has variational formulation to find $u \in V$ such that $u(X_i) = Z_i$ for $i = 1, ..., N$ and*

$$a(u,\eta) = \langle f, \eta\rangle \quad \forall \eta \in V : T\eta = 0.$$

Problem 3.3.16. *Given $f \in (H^2(\Omega))^*$, find $u^\epsilon \in V$ minimising $\frac{1}{2}a(u^\epsilon, u^\epsilon) + \frac{1}{2\epsilon}|Tu^\epsilon - Z|^2 - \langle f, u^\epsilon\rangle$. This has variational formulation to find $u^\epsilon \in V$ such that*

$$a_\epsilon(u^\epsilon, \eta) = \frac{1}{\epsilon}(Z, T\eta)_{\mathbb{R}^N} + \langle f, \eta\rangle \quad \forall \eta \in V.$$

Theorem 3.3.17. *There are unique solutions to both Problems 3.3.15 and 3.3.16.*

Proof. This is shown in [34] by making use of the Lax-Milgram theorem with the coercivity of a over V. □

For $f = 0$, these are the membrane problem studied in [34, 51]. In very much the same way as the preceding subsection, one may see that the point constraint problem can be written as the following PDE in distribution

$$\begin{aligned} \kappa\Delta^2 u - \sigma\Delta u + \sum_{i=1}^N \lambda_i \delta_{X_i} &= f \text{ in } \Omega, \\ u(X_i) &= Z_i \text{ for } i = 1, ..., N, \\ u|_{\partial\Omega} = \Delta u|_{\partial\Omega} &= 0. \end{aligned}$$

With the penalty problem having the distributional PDE,

$$\begin{aligned} \kappa\Delta^2 u^\epsilon - \sigma\Delta u^\epsilon + \frac{1}{\epsilon}\sum_{i=1}^N u^\epsilon \delta_{X_i} &= f + \frac{1}{\epsilon}\sum_{i=1}^N Z_i \delta_{X_i} \text{ in } \Omega \\ u|_{\partial\Omega} = \Delta u|_{\partial\Omega} &= 0. \end{aligned}$$

3.3.2.2 Second order splitting applied to this fourth order problem

Definition 3.3.18. *Let $\infty > q > 2 > p > 1$, then we define $X = W_0^{1,q}(\Omega)$, $Y = W_0^{1,p}(\Omega)$, $L = L^2(\Omega)$ and $S = \mathbb{R}^N$, with the operators*

$$\begin{aligned} c &: X \times X \to \mathbb{R}, \\ b &: X \times Y \to \mathbb{R}, \\ m &: L \times L \to \mathbb{R}, \end{aligned}$$

given by

$$\begin{aligned} c(u,\eta) &= \int_\Omega \left(\frac{\sigma}{\kappa} - 2\right) \nabla u \cdot \nabla \eta - u\eta, \\ b(\eta,\xi) &= \int_\Omega \nabla \eta \cdot \nabla \xi + \eta\xi, \\ m(\eta,\xi) &= \int_\Omega \eta\xi. \end{aligned}$$

This definition allows us to pose the problems for this flat case. Note the generalisation of Problems 3.3.15 and 3.3.16 by the inclusion of g in the right hand side of the second equations.

Problem 3.3.19. *Given $f \in (W^{1,q}(\Omega))^*$, $g \in (W^{1,p}(\Omega))^*$, find $(u, w, \lambda) \in X \times Y \times \mathbb{R}^N$ such that*

$$\begin{aligned} c(u,\eta) + b(\eta, w) + (T\eta, \lambda)_{\mathbb{R}^N} &= \langle f, \eta \rangle && \forall \eta \in X, \\ b(u,\xi) - m(w,\xi) &= \langle g, \xi \rangle && \forall \xi \in Y, \\ Tu &= Z. \end{aligned}$$

Problem 3.3.20. *Given $f \in (W^{1,q}(\Omega))^*$, $g \in (W^{1,p}(\Omega))^*$, find $(u^\epsilon, w^\epsilon) \in X \times Y$ such that*

$$\begin{aligned} c(u^\epsilon,\eta) + b(\eta, w^\epsilon) + \frac{1}{\epsilon}(Tu^\epsilon, T\eta)_{\mathbb{R}^N} &= \langle f, \eta \rangle + \frac{1}{\epsilon}(Z, T\eta)_{\mathbb{R}^N} && \forall \eta \in X, \\ b(u^\epsilon,\xi) - m(w^\epsilon,\xi) &= \langle g, \xi \rangle && \forall \xi \in Y. \end{aligned}$$

Checking the required assumptions, Asssumptions 3.2.5, 3.2.9, 3.2.10 and 3.2.6 hold almost identically as in Subsection 2.1 and gives the following theorem and corollary.

Theorem 3.3.21. *There are unique solutions to Problems 3.3.20 and 3.3.19.*

Corollary 3.3.22. *Let (u, w, λ) solve Problem 3.3.19 and (u^ϵ, w^ϵ) solve Problem 3.3.20. Then there is $C > 0$ such that*

$$\|u - u^\epsilon\|_X + \|w - w^\epsilon\|_L + \sqrt{\epsilon}\|w - w^\epsilon\|_Y + \sqrt{\epsilon}\|\epsilon^{-1}(Tu^\epsilon - Z) - \lambda\|_{\mathbb{R}^N} \le C\epsilon\|\lambda\|_{\mathbb{R}^N}.$$

3.4 Abstract finite element method

We now formulate and analyse an abstract finite element method to approximate the solutions to Problems 3.2.2 and 3.2.3. We formulate the method in the sense of an external approximation. This is motivated by our wish to apply the formulation to surface finite elements.

Definition 3.4.1. *For $h > 0$, let X_h, Y_h and S_h be finite dimensional normed vector spaces with lift operators*

$$l_h^X : X_h \to X, \quad l_h^Y : Y_h \to Y, \quad l_h^S : S_h \to S,$$

which are bounded, linear and injective and define $X_h^l := l_h^X(X_h)$, $Y_h^l := l_h^Y(Y_h)$, $S_h^l := l_h^S(S_h)$.

Let c_h, b_h, m_h denote bilinear forms such that

$$c_h : X_h \times X_h \to \mathbb{R}, \; b_h : X_h \times Y_h \to \mathbb{R}, \; m_h : Y_h \times Y_h \to \mathbb{R},$$

with m_h symmetric and let $(\cdot,\cdot)_{S_h}$ be an inner product on S_h. Furthermore, let $T_h : X_h \to S_h$ be linear and surjective with the compatibility condition that $T \circ l_h^X = l_h^S \circ T_h$. Finally, let $f_h \in X_h^*$, $g_h \in Y_h^*$ and $s_h \in S_h$.

The spaces X_h^l, Y_h^l and S_h^l may be identified as the spaces X_n, Y_n and S_n in Section 3.2. We will write $(\cdot)^l$ to be the lift of $(\cdot)$ with the appropriate lift map. For this section, we assume that the following approximations hold.

Assumption 3.4.2. *We assume that there is $C > 0$ and $k \in \mathbb{N}$ such that*

$$\begin{aligned}
|c(\eta_h^l, \xi_h^l) - c_h(\eta_h, \xi_h)| &\leq C h^k \|\eta_h^l\|_X \|\xi_h^l\|_X \quad \forall (\eta_h, \xi_h) \in X_h \times X_h,\\
|b(\eta_h^l, \xi_h^l) - b_h(\eta_h, \xi_h)| &\leq C h^k \|\eta_h^l\|_X \|\xi_h^l\|_Y \quad \forall (\eta_h, \xi_h) \in X_h \times Y_h,\\
|m(\eta_h^l, \xi_h^l) - m_h(\eta_h, \xi_h)| &\leq C h^k \|\eta_h^l\|_L \|\xi_h^l\|_L \quad \forall (\eta_h, \xi_h) \in Y_h \times Y_h,\\
|(z_h^l, \chi_h^l)_S - (z_h, \chi_h)_{S_h}| &\leq C h^k \|z_h^l\|_X \|\chi_h^l\|_S \quad \forall (z_h, \chi_h) \in S_h \times S_h,\\
|\langle f, \eta_h^l \rangle - \langle f_h, \eta_h \rangle| &\leq C h^k \|f\|_{X^*} \|\eta_h^l\|_X \quad \forall \eta_h \in X_h,\\
|\langle g, \xi_h^l \rangle - \langle g_h, \xi_h \rangle| &\leq C h^k \|g\|_{Y^*} \|\xi_h^l\|_Y \quad \forall \xi_h \in Y_h,\\
|(s, z_h^l)_S - (s_h, z_h)_{S_h}| &\leq C h^k \|s\|_S \|z_h^l\|_S \quad \forall z_h \in S_h.
\end{aligned}$$

The finite element approximations can now be formulated.

3.4.1 Finite element method for the Lagrange multiplier problem

For this subsection, we suppose Assumptions 3.2.5, 3.2.6, 3.2.10 and the following Assumption 3.4.3 hold true.

Assumption 3.4.3. *There is $\tilde{\zeta} > 0$ independent of h such that for any $(w_h, \chi_h) \in Y_h \times S_h$,*

$$\tilde{\zeta}(\|\chi_h^l\|_S + \|w_h^l\|_Y) \leq \sup_{(\eta_h, \xi_h) \in X_h \times Y_h} \frac{b(\eta_h^l, w_h^l) + (T\eta_h^l, \chi_h^l)_S + m(w_h^l, \xi_h^l)}{\|\eta_h^l\|_X + \|\xi_h^l\|_Y}.$$

Problem 3.4.4. *Find $(u_h, w_h, \lambda_h) \in X_h \times Y_h \times S_h$ such that*

$$\begin{aligned} c_h(u_h, \eta_h) + b_h(\eta_h, w_h) + (\lambda_h, T_h \eta_h)_{S_h} &= \langle f_h, \eta_h \rangle \quad \forall \eta_h \in X_h, \\ b_h(u_h, \xi_h) - m_h(w_h, \xi_h) &= \langle g_h, \xi_h \rangle \quad \forall \xi_h \in Y_h \\ (T_h u_h, z_h)_{S_h} &= (s_h, z_h)_{S_h} \quad \forall z_h \in S_h. \end{aligned}$$

Theorem 3.4.5. *For sufficiently small h, there exists a solution to Problem 3.4.4. Furthermore there is $C > 0$ independent of h such that*

$$\begin{aligned} \|u - u_h^l\|_X + \|w - w_h^l\|_Y + \|\lambda - \lambda_h^l\|_S \leq C \inf_{\substack{(\eta_h, \xi_h, \chi_h) \\ \in X_h \times Y_h \times S_h}} &\left(\|u - \eta_h^l\|_X + \|w - \xi_h^l\|_Y + \|\lambda - \chi_h^l\|_S \right) \\ &+ Ch^k \left(\|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S \right). \end{aligned}$$

Proof. The argument follows along the lines of [33, Theorem 3.1]. For existence and uniqueness, it is sufficient to show uniqueness for the homogeneous case, $f_h = g_h = s_h = 0$, as the system is linear and finite dimensional. In this case, we see that

$$c_h(u_h, u_h) + m_h(w_h, w_h) = 0 \text{ and } (T_h u_h, T_h u_h)_{S_h} = 0. \tag{3.12}$$

We write $G_h^l \colon Y^* \to X_h^l$ to be, for each $y \in Y^*$, $G_h^l y$ is the unique element such that

$$b(G_h^l y, \xi_h^l) = \langle y, \xi_h^l \rangle \quad \forall \xi_h^l \in Y_h^l,$$

as considered in Lemma 3.2.11. Now notice that in this homogeneous case,

$$\begin{aligned} \tilde{\beta} \|u_h^l - G_h^l m(w_h^l, \cdot)\|_X &\leq \sup_{\xi_h \in Y_h} \frac{b(u_h^l - G_h^l m(w_h^l, \cdot), \xi_h^l)}{\|\xi_h^l\|_Y} \\ &\leq \sup_{\xi_h \in Y_h} \frac{b(u_h^l, \xi_h^l) - b_h(u_h, \xi_h) + m_h(w_h, \xi_h) - m(w_h^l, \xi_h)}{\|\xi_h^l\|_Y} \\ &\leq Ch^k \|w_h^l\|_L, \end{aligned}$$

where the final line follows from Assumption 3.4.2 and that $\|u_h^l\|_X \leq C \|w_h^l\|_L$, which is a consequence of the second equation of the system with the discrete inf-sup (3.5).It follows from

(3.6),

$$
\begin{aligned}
C\|w_h^l\|_L^2 \leq & c(G_h^l m(w_h^l,\cdot),G_h^l m(w_h^l,\cdot)) + m(w_h^l,w_h^l) + \frac{1}{\epsilon_0}(TG_h^l m(w_h^l,\cdot),TG_h^l m(w_h^l,\cdot))_S \\
= & c(u_h^l,u_h^l) + m(w_h^l,w_h^l) + \frac{1}{\epsilon_0}(Tu_h^l,Tu_h^l)_S - c_h(u_h,u_h) - m_h(w_h,w_h) \\
& - \frac{1}{\epsilon_0}(T_h u_h, T_h u_h)_{S_h} + c(G_h^l m(w_h^l,\cdot),G_h^l m(w_h^l,\cdot)) - c(u_h^l,u_h^l) \\
& + \frac{1}{\epsilon_0}(TG_h^l m(w_h^l,\cdot),TG_h^l m(w_h^l,\cdot))_S - \frac{1}{\epsilon_0}(Tu_h^l,Tu_h^l)_S \\
\leq & \tilde{C}h^k\|w_h^l\|_L^2,
\end{aligned}
$$

where we have made use of the above bound, $\|u_h^l - G_h^l m(w_h^l,\cdot)\|_X \leq Ch^k\|w_h^l\|_L$, along with the approximation assumptions on the discrete bilinear forms, Assumption 3.4.2. Thus for sufficiently small h, $w_h^l = 0$, it follows that $\lambda_h = w_h = u_h = 0$, where $u_h = 0$ comes from the second equation and $\lambda_h = 0$ follows from the first equation and Assumption 3.4.3.

We have for any $\eta_h \in X_h$ and $\xi_h \in Y_h$,

$$
\begin{aligned}
\tilde{\beta}\|u_h^l - \eta_h^l\|_X \leq & \sup_{v_h \in Y_h} \frac{b(u_h^l - \eta_h^l, v_h^l)}{\|v_h^l\|_Y} \\
\leq & C\left(\|u - \eta_h^l\|_X + \|w - \xi_h^l\|_Y + \|w_h^l - \xi_h^l\|_L\right) + Ch^k\left(\|u_h^l\|_X + \|g\|_{Y^*} + \|w_h^l\|_L\right),
\end{aligned}
$$

where we have made use of

$$
\begin{aligned}
b(u_h^l - \eta_h^l, v_h^l) = & b(u - \eta_h^l, v_h^l) + b(u_h^l, v_h^l) - b_h(u_h, v_h) + \langle g_h, v_h\rangle - \langle g, v_h^l\rangle \\
& + m_h(w_h, v_h) - m(w_h^l, v_h^l) + m(w_h^l - \xi_h^l, v_h^l) - m(w - \xi_h^l, v_h^l).
\end{aligned}
$$

In a similar fashion, for any $\chi_h \in S_h$, one obtains from Assumption 3.4.3,

$$
\begin{aligned}
& \tilde{\zeta}(\|w_h^l - \xi_h^l\|_Y + \|\lambda_h^l - \chi_h^l\|_S) \\
& \qquad \leq \sup_{(x_h,v_h) \in X_h \times Y_h} \frac{b(x_h^l, w_h^l - \xi_h^l) + (\lambda_h^l - \chi_h^l, Tx_h^l)_S + m(w_h^l - \xi_h^l, v_h^l)}{\|x_h^l\|_X + \|v_h^l\|_Y} \\
& \qquad \leq C\left(\|u - \eta_h^l\|_X + \|w - \xi_h^l\|_Y + \|\lambda - \chi_h^l\|_S + \|u_h^l - \eta_h^l\|_X\right) \\
& \qquad\quad + Ch^k\left(\|u_h^l\|_X + \|w_h^l\|_Y + \|f\|_{X^*} + \|\lambda_h^l\|_S\right)
\end{aligned}
$$

where we have made use of

$$
\begin{aligned}
b(x_h^l, w_h^l - \xi_h^l) + (\lambda_h^l - \chi_h^l, Tx_h^l)_S =& b(x_h^l, w - \xi_h^l) + (\lambda - \chi_h^l, Tx_h^l)_S + b(x_h^l, w_h^l) - \langle f, x_h^l \rangle \\
&+ c(u, x_h^l) + (\lambda_h^l, Tx_h^l)_S + \langle f_h, x_h \rangle - (\lambda_h, T_h x_h)_{S_h} \\
&- b_h(x_h, w_h) - c_h(u_h, x_h) \\
\leq& C\|x_h^l\|_X \left(\|w - \xi_h^l\|_Y + \|\lambda - \chi_h^l\|_S \right) \\
&+ Ch^k \|x_h^l\|_X \left(\|w_h^l\|_Y + \|f\|_{X^*} + \|g\|_{Y^*} + \|\lambda_h^l\|_S \right) \\
&+ c(u - \eta_h^l, x_h^l) - c(u_h^l - \eta_h^l, x_h^l) + c(u_h^l, x_h^l) - c_h(u_h, x_h),
\end{aligned}
$$

and also

$$
\begin{aligned}
m(w_h^l - \xi_h^l, v_h^l) =& m(w - \xi_h^l, v_h^l) + m(w_h^l, v_h^l) - b(u, v_h^l) \\
&+ \langle g, v_h^l \rangle - m_h(w_h, v_h) + b_h(u_h, v_h) - \langle g_h, v_h \rangle.
\end{aligned}
$$

Combining these two inequalities gives

$$
\begin{aligned}
\|u_h^l - \eta_h^l\|_X +& \|w_h^l - \xi_h^l\|_Y + \|\lambda_h^l - \chi_h^l\|_S \\
&\leq C \left(\|u - \eta_h^l\|_X + \|w - \xi_h^l\|_Y + \|\lambda - \chi_h^l\|_S + \|w_h^l - \xi_h^l\|_L \right) \\
&\quad + Ch^k \left(\|u_h^l\|_X + \|w_h^l\|_Y + \|\lambda_h^l\|_S + \|f\|_{X^*} + \|g\|_{Y^*} \right).
\end{aligned}
\tag{3.13}
$$

All that is remaining, is to bound the L-norm which appears on the right hand side. We again use Lemma 3.2.12 to obtain,

$$
\begin{aligned}
C\|w_h^l - \xi_h^l\|_L^2 \leq& [c(u_h^l - \eta_h^l, u_h^l - \eta_h^l) + m(w_h^l - \xi_h^l, w_h^l - \xi_h^l) \\
&- c_h(u_h - \eta_h, u_h - \eta_h) - m_h(w_h - \xi_h, w_h - \xi_h)] \\
&+ [c_h(u_h - \eta_h, u_h - \eta_h) + m_h(w_h - \xi_h, w_h - \xi_h)] \\
&+ [c(G_h^l m(w_h^l - \xi_h^l, \cdot), G_h^l m(w_h^l - \xi_h^l, \cdot)) - c(u_h^l - \eta_h^l, u_h^l - \eta_h^l)] \\
&+ \frac{1}{\epsilon_0} (T((G_h^l m(w_h^l - \xi_h^l, \cdot)), T(G_h^l m(w_h^l - \xi_h^l, \cdot)))_S.
\end{aligned}
\tag{3.14}
$$

The first and third terms are dealt with in [33] (the first term is dealt with in their (3.5) and the third term immediately following), giving

$$
\begin{aligned}
|c(u_h^l - \eta_h^l, u_h^l - \eta_h^l) + m(w_h^l - \xi_h^l, w_h^l - \xi_h^l) -& c_h(u_h - \eta_h, u_h - \eta_h) - m_h(w_h - \xi_h, w_h - \xi_h)| \\
&\leq Ch^k (\mathcal{B}^2 + \mathcal{B}\|w_h^l - \xi_h^l\|_L + \|w_h^l - \xi_h^l\|_L^2),
\end{aligned}
$$

this is a consequence of Assumption 3.4.2, and

$$
|c(G_h^l m(w_h^l - \xi_h^l, \cdot), G_h^l m(w_h^l - \xi_h^l, \cdot)) - c(u_h^l - \eta_h^l, u_h^l - \eta_h^l)| \leq C(\mathcal{B}^2 + \mathcal{B}\|w_h^l - \xi_h^l\|_L),
$$

which follows from the discrete inf-sup, Assumption 3.2.10 and the definition of G_h^l. Where

$$\begin{aligned}\mathcal{B} :=&\|u-\eta_h^l\|_X + \|w-\xi_h^l\|_Y + \|\lambda-\chi_h^l\|_S\\ &+ h^k\left(\|g\|_{Y^*} + \|f\|_{X^*} + \|s\|_S + \|u_h^l\|_X + \|w_h^l\|_Y + \|\lambda_h^l\|_S + \|\eta_h^l\|_X + \|\xi_h^l\|_Y + \|\chi_h^l\|_S\right).\end{aligned}$$

The fourth term of (3.14) may be dealt with in the same way as the third and first part of the same equation. For the remaining term of (3.14), we calculate,

$$\begin{aligned}c_h(u_h-\eta_h, u_h-\eta_h)+&m_h(w_h-\xi_h, w_h-\xi_h)\\ =&[c_h(u_h-\eta_h, u_h-\eta_h) + b_h(u_h-\eta_h, w_h-\xi_h) + (\lambda_h-\chi_h, T_h(u_h-\eta_h))_{S_h}]\\ &-[b_h(u_h-\eta_h, w_h-\xi_h) - m_h(w_h-\xi_h, w_h-\xi_h)]\\ &-(\lambda_h-\chi_h, T_h(u_h-\eta_h))_{S_h}.\end{aligned} \tag{3.15}$$

We now split up the calculation of (3.15). For the second term of (3.15),

$$\begin{aligned}b_h(u_h-\eta_h, w_h-\xi_h)-&m_h(w_h-\xi_h, w_h-\xi_h)\\ =&b(u-\eta_h^l, w_h^l-\xi_h^l) - m(w-\xi_h^l, w-\xi_h^l) - \langle g, w_h^l-\xi_h^l\rangle + \langle g_h, w_h-\xi_h\rangle\\ &+ b(\eta_h^l, w_h^l-\xi_h^l) - m(\xi_h^l, w_h^l-\xi_h^l) - b_h(\eta_h, w_h-\xi_h) + m_h(\xi_h, w_h-\xi_h)\\ \le& C\|w_h^l-\xi_h^l\|_Y\left(\|u-\eta_h^l\|_X + \|w-\xi_h^l\|_Y + h^k\left(\|g\|_{Y^*} + \|\eta_h^l\|_X + \|\xi_h^l\|_Y\right)\right)\\ \le& C(\mathcal{B}^2 + \mathcal{B}\|w_h^l-\xi_h^l\|_L).\end{aligned}$$

With the first term of (3.15) resulting in,

$$\begin{aligned}c_h(u_h-\eta_h, u_h-\eta_h) + b_h(u_h-\eta_h, &w_h-\xi_h) + (\lambda_h-\chi_h, T_h(u_h-\eta_h))_{S_h}\\ =&c(u-\eta_h^l, u_h^l-\eta_h^l) + b(u_h^l-\eta_h^l, w-\xi_h^l) + (\lambda-\chi_h^l, T(u_h^l-\eta_h^l))_S\\ &-\langle f, u_h^l-\eta_h^l\rangle + \langle f_h, u_h-\eta_h\rangle + c(\eta_h^l, u_h^l-\eta_h^l) + b(u_h^l-\eta_h^l, \xi_h^l)\\ &+(\chi_h^l, T(u_h^l-\eta_h^l))_S - c_h(\eta_h, u_h-\eta_h) - b_h(u_h-\eta_h, \xi_h)\\ &-(\chi_h, T_h(u_h-\eta_h))_{S_h}\\ \le& C\|u_h^l-\eta_h^l\|_X\left(\|u-\eta_h^l\|_X + \|w-\xi_h^l\|_Y + \|\lambda-\chi_h^l\|_S\right)\\ &+ Ch^k\|u_h^l-\eta_h^l\|_X\left(\|\eta_h^l\|_X + \|\xi_h^l\|_Y + \|\chi_h^l\|_S + \|f\|_{X^*}\right)\\ \le& C(\mathcal{B}^2 + \mathcal{B}\|w_h^l-\xi_h^l\|_L).\end{aligned}$$

Finally, from the third term of (3.15), one has

$$\begin{aligned}(T_h(u_h-\eta_h),\lambda_h-\chi_h)_{S_h} =&(T(u-\eta_h^l),\lambda_h^l-\chi_h^l)_S+(T\eta_h^l,\lambda_h^l-\chi_h^l)_S-(s,\lambda_h^l-\chi_h^l)_S\\ &-(T_h\eta_h,\lambda_h-\chi_h)_{S_h}+(s_h,\lambda_h-\chi_h)_{S_h}\\ \le& C\|\lambda_h^l-\chi_h^l\|_S(\|u-\eta_h^l\|_X+h^k(\|s\|_S+\|\eta_h^l\|_X))\\ \le& C(\mathcal{B}^2+\mathcal{B}\|w_h^l-\xi_h^l\|_L).\end{aligned}$$

Thus for sufficiently small h, we have

$$\|w_h^l-\xi_h^l\|_L^2\le C(\mathcal{B}^2+\mathcal{B}\|w_h^l-\xi_h^l\|_L),$$

which after an application of Young's inequality, gives us the desired control on $\|w_h^l-\xi_h^l\|_L$. When putting this into (3.13) gives,

$$\|u_h^l-\eta_h^l\|_X+\|w_h^l-\xi_h^l\|_Y+\|\lambda_h^l-\chi_h^l\|_S\le C\mathcal{B}. \tag{3.16}$$

By choice of $\eta_h=\xi_h=\chi_h=0$, one has

$$\begin{aligned}\|u_h^l\|_X+\|w_h^l\|_Y+\|\lambda_h^l\|_S \le& C\left(\|u\|_X+\|w\|_Y+\|\lambda\|_S\right)\\ &+Ch^k\left(\|f\|_{X^*}+\|g\|_{Y^*}+\|s\|_S+\|u_h^l\|_X+\|w_h^l\|_Y+\|\lambda_h^l\|_S\right),\end{aligned}$$

which for sufficiently small h, and the estimates shown in Theorem 3.2.8, gives

$$\|u_h^l\|_X+\|w_h^l\|_Y+\|\lambda_h^l\|_S\le C\left(\|f\|_{X^*}+\|g\|_{Y^*}+\|s\|_S\right).$$

Using this and the triangle inequality gives

$$\begin{aligned}\|u-u_h^l\|_X+&\|w-w_h^l\|_Y+\|\lambda-\lambda_h^l\|_S\\ &\le C\left(\|u-\eta_h^l\|_X+\|w-\xi_h^l\|_Y+\|\lambda-\chi_h^l\|_S\right)\\ &+Ch^k\left(\|g\|_{Y^*}+\|f\|_{X^*}+\|s\|_S+\|\eta_h^l\|_X+\|\xi_h^l\|_Y+\|\chi_h^l\|_S\right).\end{aligned}$$

We are left to remove the isolated η_h^l, ξ_h^l and χ_h^l terms on the right hand side, this can be done by

$$\|\eta_h^l\|_X+\|\xi_h^l\|_Y+\|\chi_h^l\|_S\le\|u-\eta_h^l\|_X+\|w-\xi_h^l\|_Y+\|\lambda-\chi_h^l\|_S+\|u\|_X+\|w\|_Y+\|\lambda\|_S.$$

Thus for sufficiently small h, it holds

$$\begin{aligned}\|u-u_h^l\|_X+\|w-w_h^l\|_Y+\|\lambda-\lambda_h^l\|_S\le C\big[&\|u-\eta_h^l\|_X+\|w-\xi_h^l\|_Y+\|\lambda-\chi_h^l\|_S\\ &+h^k\left(\|f\|_{X^*}+\|g\|_{Y^*}+\|s\|_S\right)\big].\end{aligned}$$

Taking the infimum over $(\eta_h, \xi_h, \chi_h) \in X_h \times Y_h \times S_h$ gives the result. $\square$

In applications one may have interpolation operators which allow an error bound of the form of Ch^α for some $0 \leq \alpha \leq k$. The magnitude of α will depend on the regularity of the solution.

Corollary 3.4.6. *Suppose there are Banach spaces $\overline{X}$, $\overline{Y}$, $\overline{S}$ continuously embedded in X, Y, S respectively with $(u, w, \lambda) \in \overline{X} \times \overline{Y} \times \overline{S}$. Additionally assume there is $\overline{C}, \alpha > 0$ independent of h such that*

$$\inf_{(\eta_h,\xi_h,\chi_h)\in X_h\times Y_h\times S_h} \left(\|u - \eta_h^l\|_X + \|w - \xi_h^l\|_Y + \|\lambda - \chi_h^l\|_S\right) \leq \overline{C}h^\alpha(\|u\|_{\overline{X}} + \|w\|_{\overline{Y}} + \|\lambda\|_{\overline{S}}).$$

Then for sufficiently small h, there is $C > 0$ such that

$$\|u-u_h^l\|_X + \|w-w_h^l\|_Y + \|\lambda-\lambda_h^l\|_S \leq Ch^{\min(\alpha,k)}(\|u\|_{\overline{X}} + \|w\|_{\overline{Y}} + \|\lambda\|_{\overline{S}} + \|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S).$$

Proposition 3.4.7. *Under the assumptions of the above, further suppose there are Hilbert spaces H, J, K with X, Y, S continuously embedded into H, J, K respectively. Suppose there is a unique $(\psi, \phi, \omega) \in X \times Y \times S$ solving*

$$\begin{aligned} c(\eta,\psi) + b(\eta,\phi) + (T\eta,\omega)_S &= (u - u_h^l, \eta)_H && \forall \eta \in X, \\ b(\psi,\xi) - m(\phi,\xi) &= (w - w_h^l, \xi)_J && \forall \xi \in Y, \\ (T\psi, z)_S &= (\lambda - \lambda_h^l, z)_K && \forall z \in S. \end{aligned}$$

Assume there are Banach spaces $\hat{X}$, $\hat{Y}$, $\hat{S}$ continuously embedded in X, Y, S respectively, with $(\psi, \phi, \omega) \in \hat{X} \times \hat{Y} \times \hat{S}$ and there is $\hat{C}, \beta > 0$ such that

$$\inf_{(\eta_h,\xi_h,\chi_h)\in X_h\times Y_h\times S_h} \left(\|\psi - \eta_h^l\|_X + \|\phi - \xi_h^l\|_Y + \|\omega - \chi_h^l\|_S\right) \leq \hat{C}h^\beta(\|\psi\|_{\hat{X}} + \|\phi\|_{\hat{Y}} + \|\omega\|_{\hat{S}}).$$

Finally assume the regularity result of

$$\|\psi\|_{\hat{X}} + \|\phi\|_{\hat{Y}} + \|\omega\|_{\hat{S}} \leq \tilde{C}(\|u - u_h^l\|_H + \|w - w_h^l\|_J + \|\lambda - \lambda_h^l\|_K).$$

Then for sufficiently small h, there is $C > 0$ independent of h such that

$$\|u-u_h^l\|_H + \|w-w_h^l\|_J + \|\lambda-\lambda_h^l\|_K \leq Ch^{\min(\alpha+\beta,k)}(\|u\|_{\overline{X}} + \|w\|_{\overline{Y}} + \|\lambda\|_{\overline{S}} + \|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S).$$

Proof. Let (ψ, ϕ, ω) be as above, then by testing the system with $(u - u_h^l, w - w_h^l, \lambda - \lambda_h^l)$ and adding together, one has,

$$
\begin{aligned}
(u - u_h^l,&u - u_h^l)_H + (w - w_h^l, w - w_h^l)_J + (\lambda - \lambda^l, \lambda - \lambda^l)_K \\
&= c(u - u_h^l, \psi - \eta_h^l) + b(u - u_h^l, \phi - \xi_h^l) + (T(u - u_h^l), \omega - \chi_h^l)_S \\
&\quad + b(\psi - \eta_h^l, w - w_h^l) - m(\phi - \xi_h^l, w - w_h^l) + (T(\psi - \eta_h^l), \lambda - \lambda_h^l)_S \\
&\quad + \langle f, \eta_h^l\rangle + \langle g, \xi_h^l\rangle + (s, \chi_h^l)_S \\
&\quad - c(u_h^l, \eta_h^l) - b(u_h^l, \xi_h^l) - (Tu_h^l, \chi_h^l)_S - b(\eta_h^l, w_h^l) + m(\xi_h^l, w_h^l) - (T\eta_h^l, \lambda_h^l)_S \\
&\quad - \langle f_h, \eta_h\rangle - \langle g_h, \xi_h\rangle - (s_h, \chi_h)_{S_h} \\
&\quad + c_h(u_h, \eta_h) + b_h(u_h, \xi_h) + (T_h u_h, \chi_h)_{S_h} + b_h(\eta_h, w_h) - m_h(\xi_h, w_h) + (T_h\eta_h, \lambda_h)_{S_h}.
\end{aligned}
$$

we see that the first two lines may be bounded by

$$
(\|\psi - \eta_h^l\|_X + \|\phi - \xi_h^l\|_Y + \|\omega - \chi_h^l\|_S)(\|u - u_h^l\|_X + \|w - w_h^l\|_Y + \|\lambda - \lambda_h^l\|_S),
$$

and the final four lines may be bounded by the approximation properties of the discrete operators. One then obtains

$$
\begin{aligned}
(u-&u_h^l, u - u_h^l)_H + (w - w_h^l, w - w_h^l)_J + (\lambda - \lambda^l, \lambda - \lambda^l)_K \\
\leq& C\left[\left(\|\psi - \eta_h^l\|_X + \|\phi - \xi_h^l\|_Y + \|\omega - \chi_h^l\|_S\right)\left(\|u - u_h^l\|_X + \|w - w_h^l\|_Y + \|\lambda - \lambda_h^l\|_S\right)\right. \\
&+ h^k\left(\|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S\right)\left(\|\psi - \eta_h^l\|_X + \|\phi - \xi_h^l\|_Y + \|\omega - \chi_h^l\|_S\right) \\
&\quad \left. + h^k\left(\|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S\right)\left(\|\psi\|_X + \|\phi\|_Y + \|\omega\|_S\right)\right].
\end{aligned}
$$

By taking infimum over $(\eta_h, \xi_h, \chi_h) \in X_h \times Y_h \times S_h$, one has the result by use of Young's inequality. □

3.4.2 Finite element method for the penalty problem

Problem 3.4.8. *Given $\epsilon > 0$, find $(u_h^\epsilon, w_h^\epsilon) \in X_h \times Y_h$ solving the problem*

$$
\begin{aligned}
c_h(u_h^\epsilon, \eta_h) + b_h(\eta_h, w_h^\epsilon) + \frac{1}{\epsilon}(T_h u_h^\epsilon, T_h\eta_h)_{S_h} &= \langle f_h, \eta_h\rangle + \frac{1}{\epsilon}(s_h, T_h\eta_h)_{S_h} && \forall \eta_h \in X_h, \\
b_h(u_h^\epsilon, \xi_h) - m_h(w_h^\epsilon, \xi_h) &= \langle g_h, \xi_h\rangle && \forall \xi_h \in Y_h.
\end{aligned}
$$

We now prove well-posedness of this problem and give error estimates. For the error estimates, there are two obvious ways to proceed, first of all, one might wish to consider proceeding as though this is a problem independent of the hard constraint problem. An alternate approach is to use that the hard constraint problem is well approximated by the penalty problem and show a similar bound for the discrete problems, then use the error estimates for the hard constraint problem. Both of these approaches will be considered, starting with the approach estimating the difference to the hard constraint problem.

We first show the existence and uniqueness of a solution to Problem 3.4.8. We suppose that Assumptions 3.2.5, 3.2.6, 3.2.9, 3.2.10 and 3.4.3 hold true for the remainder of this subsection.

Theorem 3.4.9. *For sufficiently small h and sufficiently small ϵ, there exists a unique solution to Problem 3.4.8.*

Proof. Existence and uniqueness follows from the homogeneous case $f_h = g_h = s_h = 0$ as the system is linear and finite dimensional. In this homogeneous case we have that, by testing the first equation of the system with u_h^ϵ and the second equation of the system with w_h^ϵ and taking differences,

$$c_h(u_h^\epsilon, u_h^\epsilon) + \frac{1}{\epsilon}(T_h u_h^\epsilon, T_h u_h^\epsilon)_{S_h} + m_h(w_h^\epsilon, w_h^\epsilon) = 0. \tag{3.17}$$

It follows from Lemma 3.2.12 that

$$\begin{aligned}
C\|(w_h^\epsilon)^l\|_L^2 \leq & c(G_h^l m((w_h^\epsilon)^l, \cdot), G_h^l m((w_h^\epsilon)^l, \cdot)) \\
& + \frac{1}{\epsilon_0}(T(G_h^l m((w_h^\epsilon)^l, \cdot)), T(G_h^l m((w_h^\epsilon)^l, \cdot)))_S + m((w_h^\epsilon)^l, (w_h^\epsilon)^l) \\
\leq & c((u_h^\epsilon)^l, (u_h^\epsilon)^l) + m((w_h^\epsilon)^l, (w_h^\epsilon)^l) - c_h(u_h^\epsilon, u_h^\epsilon) - m_h(w_h^\epsilon, w_h^\epsilon) \\
& + c(G_h^l m((w_h^\epsilon)^l, \cdot), G_h^l m((w_h^\epsilon)^l, \cdot)) - c((u_h^\epsilon)^l, (u_h^\epsilon)^l) \\
& + \frac{1}{\epsilon_0}(T(G_h^l m((w_h^\epsilon)^l, \cdot)), T(G_h^l m((w_h^\epsilon)^l, \cdot)))_S - \frac{1}{\epsilon_0}(T_h u_h^\epsilon, T_h u_h^\epsilon)_{S_h} \\
\leq & \tilde{C} h^k \|(w_h^\epsilon)^l\|_L^2,
\end{aligned}$$

as in the proof of Theorem 3.4.5, where we have inequality, rather than equality for

$$c_h(u_h^\epsilon, u_h^\epsilon) + m_h(w_h^\epsilon, w_h^\epsilon) + \frac{1}{\epsilon_0}(T_h u_h^\epsilon, T_h u_h^\epsilon)_{S_h} \leq 0$$

Hence for sufficiently small h, it holds that $w_h^l = u_h^l = 0$. Thus there is a unique solution. ☐

We now wish to show a discrete version of Proposition 3.2.14 so that we may use the approximation theory from Theorem 3.4.5 to obtain uniform estimates on the solution to Problem 3.4.8.

Lemma 3.4.10. *Let $(u_h^\epsilon, w_h^\epsilon)$ solve Problem 3.4.8 and let (u_h, w_h, λ_h) solve Problem 3.4.4. Then for sufficiently small $\epsilon > 0$, there is $C > 0$ independent of ϵ and h such that,*

$$\|u_h^l - (u_h^\epsilon)^l\|_X + \|w_h^l - (w_h^\epsilon)^l\|_Y + \left\|\lambda_h^l - \epsilon^{-1} T\left((u_h^\epsilon)^l - u_h^l\right)\right\|_S \leq C\|\lambda_h^l\|_S \sqrt{\epsilon}.$$

Proof. As previously shown in the proof of Theorem 3.4.5,

$$\|G_h^l m((w_h^\epsilon)^l - w_h^l, \cdot) - \left((u_h^\epsilon)^l - u_h^l\right)\|_X \leq C h^k (\|u_h^l - (u_h^\epsilon)^l\|_X + \|w_h^l - (w_h^\epsilon)^l\|_L),$$

where we also know for h sufficiently small,

$$\|u_h^l - (u_h^\epsilon)^l\|_X \leq C\|w_h^l - (w_h^\epsilon)^l\|_L.$$

Adding $(\epsilon^{-1} - \epsilon_0^{-1})\|T((u_h^\epsilon)^l - u_h^l)\|_S^2$, to the statement of Lemma 3.2.12, with the above, yields,

$$\begin{aligned}
C\|w_h^l - (w_h^\epsilon)^l\|_L^2 + &(\epsilon^{-1} - \epsilon_0^{-1})\|T((u_h^\epsilon)^l - u_h^l)\|_S^2 \\
&\leq c(G_h^l m((w_h^\epsilon)^l - w_h^l, \cdot), G_h^l m((w_h^\epsilon)^l - w_h^l, \cdot)) + m((w_h^\epsilon)^l - w_h^l, (w_h^\epsilon)^l - w_h^l) \\
&\quad + \epsilon^{-1}(T((u_h^\epsilon)^l - u_h^l), T((u_h^\epsilon)^l - u_h^l))_S \\
&\leq c((u_h^\epsilon)^l - u_h^l, (u_h^\epsilon)^l - u_h^l) + m((w_h^\epsilon)^l - w_h^l, (w_h^\epsilon)^l - w_h^l) \\
&\quad + \epsilon^{-1}(T((u_h^\epsilon)^l - u_h^l), T((u_h^\epsilon)^l - u_h^l))_S + Ch^k\|(w_h^\epsilon)^l - w_h^l\|_L^2.
\end{aligned}$$

For sufficiently small h, we may hide the h^k terms in the left hand side, giving,

$$\begin{aligned}
C\|w_h^l - &(w_h^\epsilon)^l\|_L^2 + (\epsilon^{-1} - \epsilon_0^{-1})\|T((u_h^\epsilon)^l - u_h^l)\|_S^2 \\
&\leq c((u_h^\epsilon)^l - u_h^l, (u_h^\epsilon)^l - u_h^l) + m((w_h^\epsilon)^l - w_h^l, (w_h^\epsilon)^l - w_h^l) \\
&\quad + \epsilon^{-1}(T((u_h^\epsilon)^l - u_h^l), T((u_h^\epsilon)^l - u_h^l))_S \\
&\leq c_h(u_h^\epsilon - u_h, u_h^\epsilon - u_h) + m_h(w_h^\epsilon - w_h, w_h^\epsilon - w_h) + \epsilon^{-1}(T_h(u_h^\epsilon - u_h), T_h(u_h^\epsilon - u_h))_{S_h} \\
&\quad + c((u_h^\epsilon)^l - u_h^l, (u_h^\epsilon)^l - u_h^l) - c_h(u_h^\epsilon - u_h, u_h^\epsilon - u_h) \\
&\quad + m((w_h^\epsilon)^l - w_h^l, (w_h^\epsilon)^l - w_h^l) - m_h(w_h^\epsilon - w_h, w_h^\epsilon - w_h) \\
&\quad + \epsilon^{-1}(T((u_h^\epsilon)^l - u_h^l), T((u_h^\epsilon)^l - u_h^l))_S - \epsilon^{-1}(T_h(u_h^\epsilon - u_h), T_h(u_h^\epsilon - u_h))_{S_h},
\end{aligned}$$

where the final three lines may be bounded by $Ch^k\left(\|w_h^l - (w_h^\epsilon)^l\|_L^2 + \epsilon^{-1}\|T((u_h^\epsilon)^l - u_h^l)\|_S^2\right)$, which follows from the approximations in Assumption 3.4.2 and using the compatibility condition in Definition 3.4.1 to see that $\left(T_h(u_h^\epsilon - u_h)\right)^l = T((u_h^\epsilon)^l - u_h^l)$. Thus for h sufficiently small, one has

$$\begin{aligned}
C\big(\|w_h^l - (w_h^\epsilon)^l\|_L^2 + (\epsilon^{-1} - \epsilon_0^{-1})\|T((u_h^\epsilon)^l - u_h^l)\|_S^2\big)& \\
&\leq c_h(u_h^\epsilon - u_h, u_h^\epsilon - u_h) + m_h(w_h^\epsilon - w_h, w_h^\epsilon - w_h) \\
&\quad + \epsilon^{-1}(T_h(u_h^\epsilon - u_h), T_h(u_h^\epsilon - u_h))_{S_h} \\
&= (\lambda_h, T_h(u_h^\epsilon - u_h))_{S_h} \leq C_2\|\lambda_h^l\|_S\|T((u_h^\epsilon)^l - u_h^l)\|_S \\
&\leq C_2^2\epsilon\rho\|\lambda_h^l\|_S^2 + \frac{1}{\epsilon\rho}\|T((u_h^\epsilon)^l - u_h^l)\|_S^2,
\end{aligned}$$

where the equality follows from the discrete equations, the inequality on the penultimate line again follows from the approximation property for the inner products on S and S_h in Assumption 3.4.2 along with the compatibility condition $T \circ l_h^X = l_h^S \circ T_h$ and the final line is Young's inequality.

Thus choosing ρ sufficiently big (independent of ϵ) gives

$$\|w_h^l - (w_h^\epsilon)^l\|_L^2 + \epsilon^{-1}\|T((u_h^\epsilon)^l - u_h^l)\|_S^2 \le C\epsilon\|\lambda_h^l\|_S^2.$$

We now make use of Assumption 3.4.3,

$$\begin{aligned}
&\tilde{\zeta}(\|w_h^l - (w_h^\epsilon)^l\|_Y + \|\lambda_h^l - \epsilon^{-1}T((u_h^\epsilon)^l - u_h^l)\|_S) \\
&\qquad \le \sup_{(\eta_h,\xi_h)\in X_h\times Y_h} \frac{b(\eta_h^l, w_h^l - (w_h^\epsilon)^l) + (T\eta_h^l, \lambda_h^l - \epsilon^{-1}T((u_h^\epsilon)^l - u_h^l))_S + m(w_h^l - (w_h^\epsilon)^l, \xi_h^l)}{\|\eta_h^l\|_X + \|\xi_h^l\|_Y}.
\end{aligned}$$

It is clear in the above that the m term is bounded as we would like. We then have

$$\begin{aligned}
b(\eta_h^l, w_h^l-&(w_h^\epsilon)^l) + (T\eta_h^l, \lambda_h^l - \epsilon^{-1}T((u_h^\epsilon)^l - u_h^l))_S \\
=&b_h(\eta_h, w_h^\epsilon - w_h) + (T_h\eta_h, \lambda_h - \epsilon^{-1}T_h(u_h^\epsilon - u_h))_{S_h} + b(\eta_h^l, w_h^l - (w_h^\epsilon)^l) \\
&- b_h(\eta_h, w_h^\epsilon - w_h) + (T\eta_h^l, \lambda_h^l - \epsilon^{-1}T((u_h^\epsilon)^l - u_h^l))_S - (T_h\eta_h, \lambda_h - \epsilon^{-1}T_h(u_h^\epsilon - u_h))_{S_h} \\
=&c_h(u_h - u_h^\epsilon, \eta_h) + b(\eta_h^l, w_h^l - (w_h^\epsilon)^l) - b_h(\eta_h, w_h^\epsilon - w_h) \\
&+ (T\eta_h^l, \lambda_h^l - \epsilon^{-1}T((u_h^\epsilon)^l - u_h^l))_S - (T_h\eta_h, \lambda_h - \epsilon^{-1}T_h(u_h^\epsilon - u_h))_{S_h} \\
\le&c_h(u_h - u_h^\epsilon, \eta_h) + Ch^k\|\eta_h^l\|_X\left(\|w_h^l - (w_h^\epsilon)^l\|_Y + \|\lambda_h^l - \epsilon^{-1}T((u_h^\epsilon)^l - u_h^l)\|_S\right).
\end{aligned}$$

Thus, for h sufficiently small

$$\|w_h^l - (w_h^\epsilon)^l\|_Y + \|\lambda_h^l - \epsilon^{-1}T((u_h^\epsilon)^l - u_h^l)\|_S \le C\|u_h^l - (u_h^\epsilon)^l\|_X,$$

which completes the result. □

This results in the following Theorem.

Theorem 3.4.11. *Let $(u_h^\epsilon, w_h^\epsilon)$ be the solution to Problem 3.4.8, let (u^ϵ, w^ϵ) be the solution to Problem 3.2.3 and let (u, w, λ) be the solution to problem 3.2.2. Then there is $C > 0$ independent of h and ϵ such that*

$$\begin{aligned}
\|u^\epsilon - (u_h^\epsilon)^l\|_X + \|w^\epsilon - (w_h^\epsilon)^l\|_Y \le C &\inf_{(\eta_h,\xi_h,\chi_h)\in X_h\times Y_h\times S_h} \left(\|u - \eta_h^l\|_X + \|w - \xi_h^l\|_Y + \|\lambda - \chi_h^l\|_S\right) \\
&+ C(h^k + \sqrt{\epsilon})\left(\|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S\right).
\end{aligned}$$

Proof. We start by considering $u^\epsilon - (u_h^\epsilon)^l = (u^\epsilon - u) + (u - u_h^l) + (u_h^l - (u_h^\epsilon)^l)$ and similarly for w terms. From Proposition 3.2.14 and Lemma 3.4.10, it holds

$$\|u^\epsilon - (u_h^\epsilon)^l\|_X + \|w^\epsilon - (w_h^\epsilon)^l\|_Y \le C\sqrt{\epsilon}(\|\lambda\|_S + \|\lambda_h^l\|_S) + \|u - u_h^l\|_X + \|w - w_h^l\|_Y,$$

which combined with Theorem 3.4.5 gives the result. □

If (u^ϵ, w^ϵ) were to be sufficiently more regular than (u, w, λ), one may wish to use this extra regularity to pay for the ϵ cost and obtain higher order convergence than would be attained from Proposition 3.4.7.

3.4.3 Estimates based only on penalty formulation

Theorem 3.4.12. *Let $(u_h^\epsilon, w_h^\epsilon)$ be the solution to Problem 3.4.8, let (u^ϵ, w^ϵ) be the solution to Problem 3.2.3. Then there is $C > 0$ independent of h and ϵ such that for h^k/ϵ sufficiently small*

$$\begin{aligned}\|u^\epsilon - (u_h^\epsilon)^l\|_X + \|w^\epsilon - (w_h^\epsilon)^l\|_Y \leq & C \inf_{(\eta_h,\xi_h)\in X_h\times Y_h} \left(\left(1+\epsilon^{-1}\right) \|u^\epsilon - \eta_h^l\|_X + \|w^\epsilon - \xi_h^l\|_Y \right) \\ & + Ch^k \left(1+\epsilon^{-1}\right)^2 \left(\|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S\right).\end{aligned}$$

Proof. This argument follows as in [33, Theorem 3.1] where we keep track of the ϵ terms. For the required error estimate, we have for any $\eta_h \in X_h$ and $\xi_h \in Y_h$,

$$\begin{aligned}\tilde{\beta}\|(u_h^\epsilon)^l - \eta_h^l\|_X \leq & \sup_{v_h\in Y_h} \frac{b((u_h^\epsilon)^l - \eta_h^l, v_h^l)}{\|v_h^l\|_Y} \\ \leq & C\left(\|u^\epsilon - \eta_h^l\|_X + \|w^\epsilon - \xi_h^l\|_Y + \|(w_h^\epsilon)^l - \xi_h^l\|_L\right) \\ & + Ch^k \left(\|(u_h^\epsilon)^l\|_X + \|g\|_{Y^*} + \|(w_h^\epsilon)^l\|_L\right),\end{aligned}$$

where we have made use of

$$\begin{aligned}b((u_h^\epsilon)^l - \eta_h^l, v_h^l) =& b(u^\epsilon - \eta_h^l, v_h^l) + b((u_h^\epsilon)^l, v_h^l) - b_h(u_h^\epsilon, v_h) + \langle g_h, v_h\rangle - \langle g, v_h^l\rangle \\ & + m_h(w_h^\epsilon, v_h) - m((w_h^\epsilon)^l, v_h^l) + m((w_h^\epsilon)^l, v_h^l) - m(w^\epsilon, v_h^l).\end{aligned}$$

One then wishes to calculate $\|(w_h^\epsilon)^l - \xi_h^l\|_Y$, this is done from the inf-sup condition, giving

$$\begin{aligned}b(x_h^l, (w_h^\epsilon)^l - \xi_h^l) =& b(x_h^l, w^\epsilon - \xi_h^l) + b(x_h^l, (w_h^\epsilon)^l) + c(u^\epsilon, x_h^l) + \frac{1}{\epsilon}(Tu^\epsilon - s, Tx_h^l)_S - \langle f, x_h^l\rangle \\ & + \langle f_h, x_h\rangle - \frac{1}{\epsilon}(T_h u_h^\epsilon - s_h, T_h x_h)_{S_h} - c_h(u_h^\epsilon, x_h) - b_h(x_h, w_h^\epsilon) \\ \leq & C\|x_h^l\|_X \left(\|w^\epsilon - \xi_h^l\|_Y + \|u^\epsilon - \eta_h^l\|_X + \|(u_h^\epsilon)^l - \eta_h^l\|_X \right. \\ & \left. + h^k \left(\|f\|_{X^*} + \|(u_h^\epsilon)^l\|_X + \|(w_h^\epsilon)^l\|_Y\right)\right) \\ & + \epsilon^{-1}(Tu^\epsilon - s, Tx_h^l)_S - \epsilon^{-1}(T_h u_h^\epsilon - s_h, T_h x_h)_{S_h}.\end{aligned}$$

This yields

$$\begin{aligned}b(x_h^l, (w_h^\epsilon)^l - \xi_h^l) \leq C\|x_h^l\|_X & \left(\|w^\epsilon - \xi_h^l\|_Y + \|u^\epsilon - \eta_h^l\|_X + \|(u_h^\epsilon)^l - \eta_h^l\|_X \right. \\ & + \epsilon^{-1}\left(\|T(u^\epsilon - \eta_h^l)\|_S + \|T((u_h^\epsilon)^l - \eta_h^l)\|_S\right) \\ & \left. + h^k \left(\|f\|_{X^*} + \epsilon^{-1}\|s\|_S + \left(1+\epsilon^{-1}\right)\|(u_h^\epsilon)^l\|_X + \|(w_h^\epsilon)^l\|_Y\right)\right).\end{aligned}$$

Thus, one has

$$
\begin{aligned}
\|(u_h^\epsilon)^l - \eta_h^l\|_X + \|(w_h^\epsilon)^l - \xi_h^l\|_Y \\
\le & C[\|w^\epsilon - \xi_h^l\|_Y + \left(1+\epsilon^{-1}\right)(\|u^\epsilon - \eta_h^l\|_X + \|(w_h^\epsilon)^l - \xi_h^l\|_L) \\
& + h^k(\|f\|_{X^*} + \epsilon^{-1}\|s\|_S + \left(1+\epsilon^{-1}\right)\left(\|g\|_{Y^*} + \|(u_h^\epsilon)^l\|_X + \|(w_h^\epsilon)^l\|_L\right) + \|(w_h^\epsilon)^l\|_Y].
\end{aligned}
$$

We are now left to control $\|(w_h^\epsilon)^l - \xi_h^l\|_L$, this is done as in the proof of Theorem 3.4.5, by using the discrete coercivity relation, Lemma 3.2.12, and splitting it up into three parts. The first and third terms are identical to before, thus we are left to bound

$$
\begin{aligned}
c_h(u_h^\epsilon - \eta_h, u_h^\epsilon - \eta_h) + m_h(w_h^\epsilon - \xi_h, w_h^\epsilon - \xi_h) =& c_h(u_h^\epsilon - \eta_h, u_h^\epsilon - \eta_h) + b_h(u_h^\epsilon - \eta_h, w_h^\epsilon - \xi_h) \\
& - b_h(u_h^\epsilon - \eta_h, w_h^\epsilon - \xi_h) + m_h(w_h^\epsilon - \xi_h, w_h^\epsilon - \xi_h).
\end{aligned}
\tag{3.18}
$$

For the second term of this, we see that it is bounded by

$$
\begin{aligned}
\mathcal{B}_\epsilon :=& \|w^\epsilon - \xi_h^l\|_Y + \left(1+\epsilon^{-1}\right)\|u^\epsilon - \eta_h^l\|_X + h^k\left(\|f\|_{X^*} + \|(w_h^\epsilon)^l\|_Y + \|\xi_h^l\|_Y + \epsilon^{-1}\|s\|_S\right) \\
& + h^k\left(\left(1+\epsilon^{-1}\right)\left(\|g\|_{Y^*} + \|(u_h^\epsilon)^l\|_X + \|(w_h^\epsilon)^l\|_L + \|\eta_h^l\|_X\right)\right.
\end{aligned}
$$

For the first term of (3.18),

$$
\begin{aligned}
c_h(u_h^\epsilon - \eta_h, & u_h^\epsilon - \eta_h) + b_h(u_h^\epsilon - \eta_h, w_h^\epsilon - \xi_h) \\
\le & c_h(u_h^\epsilon - \eta_h, u_h^\epsilon - \eta_h) + b_h(u_h^\epsilon - \eta_h, w_h^\epsilon - \xi_h) + \epsilon^{-1}(T_h(u_h^\epsilon - \eta_h), T_h(u_h^\epsilon - \eta_h))_{S_h} \\
= & \epsilon^{-1}(s_h, T_h(u_h^\epsilon - \eta_h))_{S_h} + \langle f_h, u_h^\epsilon - \eta_h\rangle - c_h(\eta_h, u_h^\epsilon - \eta_h) - b_h(u_h^\epsilon - \eta_h, \xi_h) \\
& - \epsilon^{-1}(T_h\eta_h, T_h(u_h^\epsilon - \eta_h))_{S_h} - \epsilon^{-1}(s, T((u_h^\epsilon)^l - \eta_h^l))_S - \langle f, (u_h^\epsilon)^l - \eta_h^l\rangle \\
& + c(\eta_h^l, (u_h^\epsilon)^l - \eta_h^l) + b((u_h^\epsilon)^l - \eta_h^l, \xi_h^l) + \epsilon^{-1}(T\eta_h^l, T((u_h^\epsilon)^l - \eta_h^l))_S \\
& + c(u^\epsilon - \eta_h^l, (u_h^\epsilon)^l - \eta_h^l) + b((u_h^\epsilon)^l - \eta_h^l, w^\epsilon - \xi_h^l) + \epsilon^{-1}(T(u^\epsilon - \eta_h^l), T((u_h^\epsilon)^l - \eta_h^l))_S \\
\le & C\|(u_h^\epsilon)^l - \eta_h^l\|_X\left(\|u^\epsilon - \eta_h^l\|_X + \|w^\epsilon - \xi_h^l\|_Y + \epsilon^{-1}\|T(u^\epsilon - \eta_h^l)\|_S\right. \\
& \left. + h^k(\epsilon^{-1}\|s\|_S + \|f\|_{X^*} + \|\eta_h^l\|_X + \epsilon^{-1}\|T\eta_h^l\|_S) + \|\xi_h^l\|_Y\right) \\
\le & C\mathcal{B}_\epsilon(\mathcal{B}_\epsilon + \|(w_h^\epsilon)^l - \xi_h^l\|_L).
\end{aligned}
$$

Together, this yields,

$$
\|(u_h^\epsilon)^l - \eta_h^l\|_X + \|(w_h^\epsilon)^l - \xi_h^l\|_Y \le C\mathcal{B}_\epsilon
$$

A choice of $\eta_h = \xi_h = 0$ gives

$$\begin{aligned}\|(u_h^\epsilon)^l\|_X + \|(w_h^\epsilon)^l\|_Y \leq & C\left(\|w^\epsilon\|_Y + \left(1+\epsilon^{-1}\right)\|u^\epsilon\|_X + h^k\left(\|f\|_{X^*} + \|(w_h^\epsilon)^l\|_Y + \epsilon^{-1}\|s\|_S\right)\right) \\ & + Ch^k\left(1+\epsilon^{-1}\right)\left(\|g\|_{Y^*} + \|(u_h^\epsilon)^l\|_X + \|(w_h^\epsilon)^l\|_L\right),\end{aligned}$$

thus for $\frac{h^k}{\epsilon}$ sufficiently small,

$$\|(u_h^\epsilon)^l\|_X + \|(w_h^\epsilon)^l\|_Y \leq C\left(\|w^\epsilon\|_Y + \left(1+\epsilon^{-1}\right)\|u^\epsilon\|_X + h^k\left(\|f\|_{X^*} + \left(1+\epsilon^{-1}\right)\left(\|s\|_S + \|g\|_{Y^*}\right)\right)\right).$$

The triangle inequality then yields,

$$\begin{aligned}\|u^\epsilon - (u_h^\epsilon)^l\|_X + & \|w^\epsilon - (w_h^\epsilon)^l\|_Y \\ \leq & \|u^\epsilon - \eta_h^l\|_X + \|w^\epsilon - \xi_h^l\|_Y + \|(u_h^\epsilon)^l - \eta_h^l\|_X + \|(w_h^\epsilon)^l - \xi_h^l\|_Y \\ \leq & C\big[\|w^\epsilon - \xi_h^l\|_Y + \left(1+\epsilon^{-1}\right)\|u^\epsilon - \eta_h^l\|_X + h^k\left(\|f\|_{X^*} + \|(w_h^\epsilon)^l\|_Y + \|\xi_h^l\|_Y\right) \\ & + h^k\left(\epsilon^{-1}\|s\|_S + \left(1+\epsilon^{-1}\right)\left(\|g\|_{Y^*} + \|(u_h^\epsilon)^l\|_X + \|(w_h^\epsilon)^l\|_L + \|\eta_h^l\|_X\right)\right)\big] \\ \leq & C\big[\|w^\epsilon - \xi_h^l\|_Y + \left(1+\epsilon^{-1}\right)\|u^\epsilon - \eta_h^l\|_X + h^k\|\xi_h^l\|_Y \\ & + h^k\left(1+\epsilon^{-1}\right)\left(\|\eta_h^l\|_X + \|f\|_{X^*}\right) + h^k(1+\epsilon^{-1})^2\left(\|g\|_{Y^*} + \|s\|_S\right)\big]\end{aligned}$$

Further applications of the triangle inequality yield

$$\begin{aligned}\|\eta_h^l\|_X \leq & \|u^\epsilon - \eta_h^l\|_X + \|u^\epsilon\|_X \leq \|u^\epsilon - \eta_h^l\|_X + C\left(\|f\|_{X^*} + (1+\epsilon^{-1})\|g\|_{Y^*} + \epsilon^{-1}\|s\|_S\right), \\ \|\xi_h^l\|_Y \leq & \|w^\epsilon - \xi_h^l\|_Y + \|w^\epsilon\|_Y \leq \|w^\epsilon - \xi_h^l\|_Y + C(1+\epsilon^{-1})(\|f\|_{X^*} + (1+\epsilon^{-1})\|g\|_{Y^*} + \epsilon^{-1}\|s\|_S).\end{aligned}$$

Thus

$$\begin{aligned}\|u^\epsilon - (u_h^\epsilon)^l\|_X + \|w^\epsilon - (w_h^\epsilon)^l\|_Y \leq C\big[& \|w^\epsilon - \xi_h^l\|_Y + \left(1+\epsilon^{-1}\right)\|u^\epsilon - \eta_h^l\|_X \\ & + h^k(1+\epsilon^{-1})\|f\|_{X^*} + h^k\left(1+\epsilon^{-1}\right)^2\left(\|g\|_{Y^*} + \|s\|_S\right)\big].\end{aligned}$$

Taking infimum over $(\eta_h, \xi_h) \in X \times Y$ gives the result. $\square$

Corollary 3.4.13. *Suppose there are Banach spaces $\overline{X}$ and $\overline{Y}$ continuously embedded in X and Y, respectively, with $(u^\epsilon, w^\epsilon) \in \overline{X} \times \overline{Y}$. Additionally assume there is $\overline{C}, \alpha > 0$ independent of h such that*

$$\inf_{(\eta_h,\xi_h)\in X_h\times Y_h}\left(\|u^\epsilon - \eta_h^l\|_X + \|w^\epsilon - \xi_h^l\|_Y\right) \leq \overline{C}h^\alpha\left(\|u^\epsilon\|_{\overline{X}} + \|w^\epsilon\|_{\overline{Y}}\right).$$

Then for sufficiently small h, there is $C > 0$ such that

$$\begin{aligned}\|u^\epsilon - (u_h^\epsilon)^l\|_X + & \|w^\epsilon - (w_h^\epsilon)^l\|_Y \\ \leq & C\left(1+\epsilon^{-1}\right)^2 h^{\min(\alpha,k)}\left(\|u^\epsilon\|_{\overline{X}} + \|w^\epsilon\|_{\overline{Y}} + \|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S\right).\end{aligned}$$

Proposition 3.4.14. *Under the assumptions of the above, further suppose there are Hilbert spaces H and J such that X and Y are continuously embedded into H and J respectively. Let $(\psi, \phi) \in X \times Y$ be the unique solution to*

$$\begin{aligned} c(\eta,\psi) + b(\eta,\phi) + \frac{1}{\epsilon}(T\eta, T\psi)_S &= (u^\epsilon - (u_h^\epsilon)^l, \eta)_H \quad \forall \eta \in X \\ b(\psi,\xi) - m(\phi,\xi) &= (w^\epsilon - (w_h^\epsilon)^l, \xi)_J \quad \forall \xi \in Y. \end{aligned}$$

Assume there are Banach spaces $\hat{X}$, $\hat{Y}$ continuously embedded in X, Y respectively, with $(\psi,\phi) \in \hat{X} \times \hat{Y}$ and there is $\hat{C}, \beta > 0$ such that

$$\inf_{(\eta_h,\xi_h) \in X_h \times Y_h} \left(\|\psi - \eta_h^l\|_X + \|\phi - \xi_h^l\|_Y \right) \leq \hat{C} h^\beta \left(\|\psi\|_{\hat{X}} + \|\phi\|_{\hat{Y}} \right).$$

Finally assume the regularity result of

$$\|\psi\|_{\hat{X}} + \|\phi\|_{\hat{Y}} \leq \tilde{C}(\|u^\epsilon - (u_h^\epsilon)^l\|_H + \|w^\epsilon - (w_h^\epsilon)^l\|_J).$$

Then for sufficiently small h, there is $C > 0$ independent of h such that

$$\begin{aligned} \|u^\epsilon - (u_h^\epsilon)^l\|_H + &\|w^\epsilon - (w_h^\epsilon)^l\|_J \\ &\leq C h^{\min(\alpha+\beta,k)} \left(1 + \epsilon^{-1}\right)^3 \left(\|u^\epsilon\|_{\overline{X}} + \|w^\epsilon\|_{\overline{Y}} + \|f\|_{X^*} + \|g\|_{Y^*} + \|s\|_S\right). \end{aligned}$$

Proof. Let (ψ, ϕ) be as above, then by testing with $(u^\epsilon - (u_h^\epsilon)^l, w^\epsilon - (w_h^\epsilon)^l)$ and summing, one has

$$\begin{aligned} (u^\epsilon - (u_h^\epsilon)^l, &u^\epsilon - (u_h^\epsilon)^l)_H + (w^\epsilon - (w^\epsilon)^l, w^\epsilon - (w_h^\epsilon)^l)_J \\ &= c(u^\epsilon - (u_h^\epsilon)^l, \psi - \eta_h^l) + b(u^\epsilon - (u_h^\epsilon)^l, \phi - \xi_h^l) + \epsilon^{-1}(T(u^\epsilon - (u_h^\epsilon)^l), T(\psi - \eta_h^l))_S \\ &\quad + b(\psi - \eta_h^l, w^\epsilon - (w_h^\epsilon)^l) - m(\phi - \xi_h^l, w^\epsilon - (w_h^\epsilon)^l) \\ &\quad + \langle f, \eta_h^l \rangle + \epsilon^{-1}(s, T\eta_h^l)_S + \langle g, \xi_h^l \rangle \\ &\quad - c((u_h^\epsilon)^l, \eta_h^l) - b((u_h^\epsilon)^l, \xi_h^l) - \epsilon^{-1}(T(u_h^\epsilon)^l, T\eta_h^l)_S - b(\eta_h^l, (w_h^\epsilon)^l) + m(\xi_h^l, (w_h^\epsilon)^l) \\ &\quad - \langle f_h, \eta_h \rangle - \epsilon^{-1}(s_h, T_h\eta_h)_{S_h} - \langle g_h, \xi_h \rangle \\ &\quad + c_h(u_h^\epsilon, \eta_h) + b_h(u_h^\epsilon, \xi_h) + \epsilon^{-1}(T_h u_h^\epsilon, \eta_h)_{S_h} + b_h(\eta_h, w_h^\epsilon) - m_h(\xi_h, w_h^\epsilon), \end{aligned}$$

we see the first two lines may be bounded by

$$((1 + \epsilon^{-1})\|\psi - \eta_h^l\|_X + \|\phi - \xi_h^l\|_Y)(\|u^\epsilon - (u_h^\epsilon)^l\|_X + \|w^\epsilon - (w_h^\epsilon)^l\|_Y),$$

with the final four lines being bounded by the approximating properties of the discrete bilinear forms and linear functionals. It follows,

$$
\begin{aligned}
(u^\epsilon-&(u_h^\epsilon)^l, u^\epsilon-(u_h^\epsilon)^l)_H + (w^\epsilon-(w^\epsilon)^l, w^\epsilon-(w_h^\epsilon)^l)_J \\
\leq & C\left[\left(\left(1+\epsilon^{-1}\right)\|\psi-\eta_h^l\|_X + \|\phi-\xi_h^l\|_Y\right)\left(\|u^\epsilon-(u_h^\epsilon)^l\|_X + \|w^\epsilon-(w_h^\epsilon)^l\|_Y\right)\right. \\
& \left. + h^k\left(1+\epsilon^{-1}\right)(\|f\|_{X^*}+\|g\|_{Y^*}+\|s\|_S)\left(\|\psi-\eta_h^l\|_X + \|\phi-\xi_h^l\|_Y + \|\psi\|_X + \|\phi\|_Y\right)\right].
\end{aligned}
$$

Taking infimum over $(\eta_h, \xi_h) \in X_h \times Y_h$ and applying Young's inequality completes the result. □

3.5 Finite element approximation for PDE example problems

3.5.1 A near spherical biomembrane

Definition 3.5.1. *Define the following bilinear forms on the discrete function space*

$$
\begin{aligned}
c_h(u_h, \eta_h) &= \int_{\Gamma_h} \left(\frac{\sigma}{\kappa} - 2 - \frac{2}{R^2}\right)\nabla_{\Gamma_h} u_h \cdot \nabla_{\Gamma_h}\eta_h - \left(1 + \frac{2\sigma}{\kappa R^2}\right) u_h \eta_h, \\
b_h(\eta_h, \xi_h) &= \int_{\Gamma_h} \nabla_{\Gamma_h}\eta_h \cdot \nabla_{\Gamma_h}\xi + \eta_h \xi_h, \\
m_h(\eta_h, \xi_h) &= \int_{\Gamma_h} \eta_h \xi_h,
\end{aligned}
$$

with $T_h\eta_h := T\eta_h^l$. We take $f_h, g_h \in (\mathcal{S}_h^l)^*$ to satisfy $\langle f_h, \eta_h\rangle = \langle f, \eta_h^l\rangle$, $\langle g_h, \xi_h\rangle = \langle g, \xi_h^l\rangle$.

We now verify Assumption 3.4.3.

Proposition 3.5.2. *There is $C > 0$ such that*

$$
C(\|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N}) \leq \sup_{(\eta_h^l, \xi_h^l)\in(\mathcal{S}_h^l\cap X)\times(\mathcal{S}_h^l\cap Y))} \frac{b(\eta_h^l, w_h^l) + (\lambda_h, T\eta_h^l) + m(w_h^l, \xi_h^l))}{\|\eta_h^l\|_{1,q} + \|\xi_h^l\|_{1,p}}
$$

Proof. Well-posedness of Problem 3.3.8 gives that for any $(u, w, \lambda) \in X \times Y \times \mathbb{R}^N$, there is $C > 0$ such that

$$
\begin{aligned}
C(\|u\|_{1,q}+&\|w\|_{1,p} + \|\lambda\|_{\mathbb{R}^N}) \\
&\leq \sup_{(\eta,\xi,\chi)\in X\times Y\times\mathbb{R}^N} \frac{c(u,\eta) + b(\eta, w) + (T\eta, \lambda) + b(u, \xi) - m(w, \xi) + (Tu, \chi)_{\mathbb{R}^N}}{\|\eta\|_{1,q} + \|\xi\|_{1,p} + \|\chi\|_{\mathbb{R}^N}}.
\end{aligned}
$$

By applying this with $(0, y_h^l, \chi_h) \in X \times Y \cap \mathcal{S}_h^l \times \mathbb{R}^N$,

$$
\begin{aligned}
\|y_h^l\|_{1,p} + &\|\chi_h\|_{\mathbb{R}^N} \\
\leq & C \sup_{(\eta,\xi)\in X\times Y} \frac{b(\eta, y_h^l) + (T\eta, \chi_h) + m(y_h^l, \xi)}{\|\eta\|_{1,q} + \|\xi\|_{1,p}} \\
= & C \sup_{(\eta,\xi)\in X\times Y} \frac{b(\Pi_h\eta, y_h^l) + (T\Pi_h\eta, \chi_h) + (T(\eta - \Pi_h\eta), \chi_h) + m(y_h^l, P_h\xi)}{\|\eta\|_{1,q} + \|\xi\|_{1,p}} \\
\leq & C \sup_{(\eta,\xi)\in X\times Y} \frac{b(\Pi_h\eta, y_h^l) + (T\Pi_h\eta, \chi_h) + m(y_h^l, P_h\xi)}{\|\Pi_h\eta\|_{1,q} + \|P_h\xi\|_{1,p}} + C' h^{1-2/q} |\log(h)| \|\chi_h\|_{\mathbb{R}^N}.
\end{aligned}
$$

Where P_h is the $L^2(\Gamma)$ projection, the log term appears from $\|\eta - \Pi_h\eta\|_{0,\infty} \leq C|\log(h)|\|\eta - I_h^l\eta\|_{0,\infty}$ [74], the $h^{1-2/q}$ follows from interpolation inequalities and P_h is a bounded operator from $W^{1,p}(\Gamma)$ to itself [11]. Thus for sufficiently small h, this completes the proof. $\square$

Problem 3.5.3. *Find $(u_h, w_h, \lambda_h) \in \mathcal{S}_h \times \mathcal{S}_h \times \mathbb{R}^N$ such that $\int_{\Gamma_h} u_h = 0$ and*

$$
\begin{aligned}
c_h(u_h, \eta_h) + b_h(\eta_h, w_h) + (T_h\eta_h, \lambda_h)_{\mathbb{R}^N} &= \langle f_h, \eta_h \rangle \quad \forall \eta_h \in \mathcal{S}_h : \int_{\Gamma_h} \eta_h = 0, \\
b_h(u_h, \xi_h) - m_h(w_h, \eta_h) &= \langle g_h, \xi_h \rangle \quad \forall \xi_h \in \mathcal{S}_h, \\
T_h u_h &= Z.
\end{aligned}
$$

Theorem 3.5.4. *There is a unique solution to Problem 3.5.3. Moreover, for $g \in L^2(\Gamma)$ it holds that*

$$
\|u - u_h^l\|_{1,2} + \|w - w_h^l\|_{0,2} + \|\lambda - \lambda_h^l\|_{\mathbb{R}^N} \leq C h^{2/q} (\|f\|_{-1,p} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N}).
$$

One might hope that the estimate follows from Theorem 3.4.5, Corollary 3.4.6 and Proposition 3.4.7. However, it is possible to see that due to our choice of T the maximum regularity one might expect is $\overline{X} = H^2(\Gamma)$, $\hat{X} = W^{3,p}(\Gamma)$, $\overline{Y} = \hat{Y} = W^{1,p}(\Gamma)$, and $\overline{S} = \hat{S} = S$, which would give $\alpha = \beta = 0$ in the context of Proposition 3.4.7. As such we require a different method, the idea is to, in the proof of Proposition 3.4.7, pick ξ_h to be $\Pi_h w$ which gives that the term which would depend on $\|\phi - \xi_h^l\|_Y$ vanishes. We also address the fact that the typical lift map from the discrete surface to the continuous surface will not, in general, preserve the integral of functions.

Proof. The existence follows from Theorem 3.4.5. For the estimate, consider $(\psi, \phi, \chi) \in X \times W^{1,p}(\Gamma) \times \mathbb{R}^N$ such that

$$
\begin{aligned}
c(\eta, \psi) + b(\eta, \phi) + (T\eta, \chi) &= (u - u_h^l, \eta)_{H^1(\Gamma)} \quad \forall \eta \in X, \\
b(\psi, \xi) - m(\phi, \xi) &= (w - w_h^l, \xi)_{L^2(\Gamma)} \quad \forall \xi \in W^{1,p}(\Gamma), \\
T\psi &= \lambda - \lambda_h.
\end{aligned}
$$

This has a unique solution with $\|\psi\|_{2,2} + \|\phi\|_{1,p} + \|\chi\|_{\mathbb{R}^N} \leq C(\|u - u_h^l\|_{1,2} + \|w - w_h^l\|_{0,2} + \|\lambda -$

$\lambda_h\|_{\mathbb{R}^N})$. Testing this system with $(u - u_h^l + [u_h^l], w - w_h^l, \lambda - \lambda_h)$, where $[v] := \frac{1}{|\Gamma|}\int_\Gamma v$ is the average value of v, gives

$$\begin{aligned}\|u - u_h^l\|_{1,2}^2 + (u - u_h^l, [u_h^l])_{H^1(\Gamma)} + &\|w - w_h^l\|_{0,2}^2 + \|\lambda - \lambda_h\|_{\mathbb{R}^N}^2 \\ =& c(u - u_h^l + [u_h^l], \psi) + b(\psi, w - w_h^l) + (T\psi, \lambda - \lambda_h)_{\mathbb{R}^N} \\ &+ b(u - u_h^l + [u_h^l], \phi) - m(w - w_h^l, \phi) + (T(u - u_h^l), \chi)_{\mathbb{R}^N}.\end{aligned}$$

The final term here is 0 when h is sufficiently small that $T_h u_h = T u_h^l = T u = Z$, and it holds that $|[u_h^l]| \leq C h^2 \|u_h^l\|_{0,2}$ since

$$\left|\int_\Gamma u_h^l\right| = \left|\int_\Gamma u_h^l - \int_{\Gamma_h} u_h\right| \leq c h^2 \|u_h^l\|_{0,2} \|1\|_{0,2}.$$

We consider

$$\begin{aligned}b(u - u_h^l, \phi) - m(w - w_h^l, \phi) =& \langle g, \phi - \Pi_h \phi\rangle + \langle g, \Pi_h \phi\rangle - \langle g_h, \Pi_h \phi^{-l}\rangle \\ &+ m(w_h^l, \phi - \Pi_h\phi) + b_h(u_h, \Pi_h \phi^{-l}) - b(u_h^l, \Pi_h \phi) \\ &+ m(w_h^l, \Pi_h\phi) - m_h(w_h, \Pi_h \phi^{-l}) \\ \leq& C\big[\left(\|g\|_{0,2} + \|w_h^l\|_{0,2}\right) \|\phi - \Pi_h\phi\|_{0,2} \\ &+ h^2 \left(\|g\|_{0,2} + \|u_h^l\|_{1,q} + \|w_h^l\|_{0,2}\right) \|\phi\|_{1,p}\big] \\ \leq& C\Big[h^{2/q}(\|g\|_{0,2} + \|w_h^l\|_{0,2}) \\ &+ h^2(\|g\|_{0,2} + \|u_h^l\|_{1,q} + \|w_h^l\|_{0,2})\Big] \|\phi\|_{1,p}.\end{aligned}$$

For the remaining terms,

$$\begin{aligned}&c(u - u_h^l, \psi) + b(\psi, w - w_h^l) + (T\psi, \lambda - \lambda_h)_{\mathbb{R}^N} \\ &\quad= \langle f, \psi\rangle - c(u_h^l, \psi - I_h^l\psi) - b(\psi - I_h^l \psi, w_h^l) - (T(\psi - I_h^l\psi), \lambda_h)_{\mathbb{R}^N} \\ &\qquad - c(u_h^l, I_h^l\psi) - b(I_h^l\psi, w_h^l) - (T I_h^l \psi, \lambda_h)_{\mathbb{R}^N} - \langle f_h, I_h\psi - [I_h\psi]\rangle \\ &\qquad + c_h(u_h, I_h\psi - [I_h\psi]) + b_h(I_h\psi - [I_h\psi], w_h) + (T_h(I_h\psi - [I_h\psi]), \lambda_h)_{\mathbb{R}^N} \\ &\quad\leq C\big[(\|f\|_{-1,p} + \|u_h^l\|_{1,q} + \|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N})\|\psi - I_h^l\psi\|_{1,q} \\ &\qquad + h^2 \|I_h^l\psi\|_{1,q}(\|u_h^l\|_{1,q} + \|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N} + \|f\|_{-1,p}) \\ &\qquad + [I_h\psi](\|u_h^l\|_{1,q} + \|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N} + \|f\|_{-1,p})\big] \\ &\quad\leq C\big[h^{2/q}(\|f\|_{-1,p} + \|u_h^l\|_{1,q} + \|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N})\|\psi\|_{2,2} \\ &\qquad + h^2 \|I_h^l\psi\|_{1,q}(\|u_h^l\|_{1,q} + \|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N} + \|f\|_{-1,p}) \\ &\qquad + [I_h\psi](\|u_h^l\|_{1,q} + \|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N} + \|f\|_{-1,p})\big],\end{aligned}$$

where we may see that $[I_h\psi] \le Ch^2\|\psi\|_{2,2}$, and we have used

$$\langle f, I_h^l\psi\rangle - \langle f_h, I_h\psi\rangle \le Ch^2\|f\|_{-1,p}\|I_h^l\psi\|_{1,q},$$

which follows from the estimates on m and m_h and using density. Hence we have, after using Young's inequality on the additional left hand side term,

$$\begin{aligned}\|u - u_h^l\|_{1,2}^2 + &\|w - w_h^l\|_{0,2}^2 + \|\lambda - \lambda_h\|_{\mathbb{R}^N}^2 \\ &\le Ch^{2/q}(\|g\|_{0,2} + \|f\|_{-1,p} + \|u_h^l\|_{1,q} + \|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N})(\|\psi\|_{2,2} + \|\phi\|_{1,p}),\end{aligned}$$

which after using $\|u_h^l\|_{1,q} + \|w_h^l\|_{1,p} + \|\lambda_h\|_{\mathbb{R}^N} \le C(\|g\|_{0,2} + \|f\|_{-1,p} + \|Z\|_{\mathbb{R}^N})$, and the regularity estimates assumed for $\|\psi\|_{2,2} + \|\phi\|_{1,p}$, gives the result. □

We now wish to improve this estimates to the spaces which are natural to the problem.

Corollary 3.5.5. *Under the assumptions of the above theorem, it holds that*

$$\|u - u_h^l\|_{1,q} \le Ch^{2/q}(\|g\|_{0,2} + \|f\|_{-1,p} + \|Z\|_{\mathbb{R}^N}).$$

Proof. The inf-sup condition in Lemma 3.3.2 gives,

$$\beta\|u - u_h^l\|_{1,q} \le \sup_{\xi\in W^{1,p}(\Gamma)} \frac{b(u - u_h^l, \xi)}{\|\xi\|_{1,p}},$$

where we see

$$\begin{aligned}b(u - u_h^l, \xi) =& m(w, \xi - \Pi_h\xi) + \langle g, \xi - \Pi_h\xi\rangle + m(w - w_h^l, \Pi_h\xi) \\ &+ m(w_h^l, \Pi_h\xi) - b(u_h^l, \Pi_h\xi) + \langle g, \Pi_h\xi\rangle \\ &- m_h(w_h, (\Pi_h\xi)^{-l}) + b_h(u_h, (\Pi_h\xi)^{-l}) - \langle g_h, (\Pi_h\xi)^{-l}\rangle \\ \le& Ch^{2/q}(\|w\|_{0,2} + \|g\|_{0,2})\|\xi\|_{1,p} + C\|w - w_h^l\|_{0,2}\|\xi\|_{1,p}.\end{aligned}$$

The estimate shown for $\|w - w_h^l\|_{0,2}$ in Theorem 3.5.4 completes the result. □

Corollary 3.5.6. *In addition to the assumptions of the above, assume $f \in C(\Gamma)^*$, it then holds,*

$$\|w - w_h^l\|_{1,p} \le Ch^{\min(2/p-1,2/q)}|\log(h)|(\|f\|_{C(\Gamma)^*} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N}).$$

Proof. The inf-sup condition from Proposition 3.3.5 gives,

$$\overline{\gamma}\|w - w_h^l - [w] + [w_h^l]\|_{1,p} \le \sup_{\eta\in X} \frac{b(\eta, w - w_h^l - [w] + [w_h^l])}{\|\eta\|_{1,q}} = \sup_{\eta\in X} \frac{b(\eta, w - w_h^l)}{\|\eta\|_{1,q}}.$$

Thus we calculate

$$
\begin{aligned}
b(\eta, w - w_h^l) =& \langle f, \eta - \Pi_h \eta \rangle - c(u, \eta - \Pi_h \eta) - (\lambda, T(\eta - \Pi_h \eta))_{\mathbb{R}^N} \\
&+ \langle f, \Pi_h \eta \rangle + c(u_h^l - u, \Pi_h \eta) + (\lambda_h - \lambda, T \Pi_h \eta)_{\mathbb{R}^N} - (\lambda_h, T \Pi_h \eta)_{\mathbb{R}^N} \\
&- c(u_h^l, \Pi_h \eta) - b(\Pi_h \eta, w_h^l) \\
&+ c_h(u_h, \Pi_h^{-l} \eta) + b_h(\Pi_h^{-l} \eta, w_h) + (\lambda_h, T_h \Pi_h^{-l} \eta)_{\mathbb{R}^N} - \langle f_h, \Pi_h^{-l} \eta \rangle.
\end{aligned}
$$

For our particular choice of c, it holds that $c(u, \eta) \leq C \|u\|_{2,2} \|\eta\|_{0,2}$. For $f \in C(\Gamma)^*$ it holds $\langle f, \eta \rangle \leq \|f\|_{C(\Gamma)^*} \|\eta\|_{0,\infty}$. Furthermore, we have that $T_h \Pi_h^{-l} \eta = T \Pi_h \eta$ by definition. From [74] $\|\eta - \Pi_h \eta\|_{0,\infty} \leq C |\log(h)| \|\eta - I_h^l \eta\|_{0,\infty}$. Together, this gives

$$
\begin{aligned}
b(\eta, w - w_h^l) \leq C \|\eta\|_{1,q} \big[& h \|u\|_{2,2} + |\log(h)| h^{1-2/q} (\|\lambda\|_{\mathbb{R}^N} + \|f\|_{C(\Gamma)^*}) \\
&+ \|u - u_h^l\|_{1,2} + \|\lambda - \lambda_h\|_{\mathbb{R}^N} + h^2 (\|f\|_{C(\Gamma)^*} + \|u_h^l\|_{1,q} + \|w_h^l\|_{1,p}) \big].
\end{aligned}
$$

Using the estimate in Theorem 3.5.4 for $\|u - u_h^l\|_{1,2} + \|\lambda - \lambda_h\|_{\mathbb{R}^N}$ completes the proof. □

Notice that for $p = 4/3$, this results in almost $h^{\frac{1}{2}}$ convergence. We now look at the problem with penalty, which follows as in Theorem 3.4.11.

Problem 3.5.7. *Find* $(u_h^\epsilon, w_h^\epsilon) \in \mathcal{S}_h^l \times \mathcal{S}_h^l$ *such that* $\int_{\Gamma_h} u_h^\epsilon = 0$ *and*

$$
\begin{aligned}
c_h(u_h^\epsilon, \eta_h) + b_h(\eta_h, w_h^\epsilon) + \frac{1}{\epsilon} (T_h u_h^\epsilon, T \eta_h)_{\mathbb{R}^N} &= \langle f_h, \eta_h \rangle + \frac{1}{\epsilon} (Z, T \eta_h)_{\mathbb{R}^N} \qquad \forall \eta_h \in \mathcal{S}_h : \int_{\Gamma_h} \eta_h = 0, \\
b_h(u_h^\epsilon, \xi_h) - m(w_h^\epsilon, \xi_h) &= \langle g_h, \xi_h \rangle \qquad \forall \xi_h \in \mathcal{S}_h.
\end{aligned}
$$

Theorem 3.5.8. *There is a unique solution to Problem 3.5.7. Moreover, for* $g \in L^2(\Gamma)$*, it holds for* $f \in W^{1,q}(\Gamma)^*$*,*

$$
\|u^\epsilon - (u_h^\epsilon)^l\|_{1,2} + \|w^\epsilon - (w_h^\epsilon)^l\|_{0,2} \leq C (h^{2/q} + \sqrt{\epsilon}) (\|f\|_{-1,p} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N})
$$

In particular,

$$
\|u^\epsilon - (u_h^\epsilon)^l\|_{1,q} \leq C (h^{2/q} + \sqrt{\epsilon}) (\|f\|_{-1,p} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N})
$$

and if $f \in C(\Gamma)^*$*, then for any* $1 < p < 2 < q < \infty$ *with* p, q *conjugate*

$$
\|w^\epsilon - (w_h^\epsilon)^l\|_{1,p} \leq C (h^{\min(2/p-1, 2/q)} |\log(h)| + \sqrt{\epsilon}) (\|f\|_{C(\Gamma)^*} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N}).
$$

Proof. The results follow from the same argument as Theorem 3.4.11, where we have the results of Theorem 3.5.4 and Corollaries 3.5.5 and 3.5.6 to give the h estimates and we may see from Corollary 3.2.17 and Lemma 3.4.10 the ϵ approximation. □

3.5.2 A flat biomembrane

We here outline the existence results and estimates as in the preceding subsection for the case of the nearly flat membrane problem discussed in Subsection 3.5.2. We make the simplifying assumption that Ω is convex and polygonal.

Definition 3.5.9. *Let $\mathcal{T}_h$ be a triangulation of Ω with $\overline{\Omega} = \bigcup_{K\in\mathcal{T}_h} K$ and $K^\circ \cap (K')^\circ = \emptyset$ $\forall K, K' \in \mathcal{T}_h$ for $K \neq K'$. Define*

$$\mathcal{S}_h := \{\chi \in C(\Omega) \,:\, \chi|_K \in P^1(K)\ \forall K \in \mathcal{T}_h\}.$$

Where $P^1(K)$ is the polynomials of degree 1 or less on K. The Lagrange basis functions ϕ_i of this space are uniquely determined by their values at the so-called Lagrange nodes q_j. The associated Lagrange interpolation $I_h\colon C(\Omega) \to \mathcal{S}_h$ is given by $I_h f := \sum_i f(q_i)\phi_i$. We again take the linear functionals as in Definition 3.3.18 and assume that $\mathcal{T}_h$ is a regular triangulation.

Problem 3.5.10. *Find $(u_h, w_h, \lambda_h) \in \mathcal{S}_h \times \mathcal{S}_h \times \mathbb{R}^N$ such that*

$$\begin{aligned} c(u_h, \eta_h) + b(\eta_h, w_h) + (T\eta_h, \lambda_h)_{\mathbb{R}^N} =& \langle f_h, \eta_h\rangle\ \forall \eta_h \in \mathcal{S}_h \\ b(u_h, \xi_h) - m(w_h, \xi_h) =& \langle g_h, \xi_h\rangle\ \forall \xi_h \in \mathcal{S}_h \\ Tu_h =& Z. \end{aligned}$$

Problem 3.5.11. *Find $(u_h^\epsilon, w_h^\epsilon) \in \mathcal{S}_h \times \mathcal{S}_h$ such that*

$$\begin{aligned} c(u_h^\epsilon, \eta_h) + b(\eta_h, w_h^\epsilon) + \frac{1}{\epsilon}(Tu_h^\epsilon, T\eta_h)_{\mathbb{R}^N} =& \langle f, \eta_h\rangle + \frac{1}{\epsilon}(Z, T\eta_h)_{\mathbb{R}^N} \quad \forall \eta_h \in \mathcal{S}_h \\ b(u_h, \xi_h) - m(w_h, \xi_h) =& \langle g, \xi_h\rangle\ \forall \xi_h \in \mathcal{S}_h \end{aligned}$$

Theorem 3.5.12. *There is a unique solution to Problem 3.5.10. Moreover for $g \in L^2(\Omega)$ and any $q > 2$, it holds for $f \in (W^{1,q}(\Omega))^*$*

$$\|u - u_h\|_{1,q} + \|w - w_h\|_{0,2} \leq C(h^{2/q}(\|f\|_{-1,p} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N})$$

Furthermore, if $f \in (C(\Omega))^*$, with $\frac{1}{p} + \frac{1}{q} = 1$,

$$\|w - w_h\|_{1,p} \leq C(h^{\min(2/p-1,2/q)}|\log(h)|)(\|f\|_{C(\Omega)^*} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N}).$$

Theorem 3.5.13. *There is a unique solution to Problem 3.5.11. Moreover for $g \in L^2(\Omega)$ and any $q > 2$, it holds for $f \in (W^{1,q}(\Omega))^*$*

$$\|u^\epsilon - u_h^\epsilon\|_{1,q} + \|w^\epsilon - w_h^\epsilon\|_{0,2} \leq C(h^{2/q} + \sqrt{\epsilon})(\|f\|_{-1,p} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N})$$

Furthermore, if $f \in (C(\Omega))^*$, *with* $\frac{1}{p} + \frac{1}{q} = 1$,

$$\|w^\epsilon - w_h^\epsilon\|_{1,p} \le C(h^{\min(2/p-1,2/q)}|\log(h)| + \sqrt{\epsilon})(\|f\|_{C(\Omega)^*} + \|g\|_{0,2} + \|Z\|_{\mathbb{R}^N}).$$

These results follow from a slight variation of the arguments presented in Subsection 3.5.1.

3.6 Numerical experiments

We conclude with some numerical examples. All of the numerical examples are done for the biomembrane problem as outlined in Section 3.3. When we discuss the error at level h, we will be referring to the relative error, where we define the relative error between u and u_h^l in norm $\|\cdot\|_W$ to be given by $E_W(h) := \|u - u_h^l\|_W / \|u\|_W$. The EOC (experimental order of convergence) between levels h_1 and h_2 is given by $EOC_W(h_1, h_2) := \log(E_W(h_1)/E_W(h_2))/\log(h_1/h_2)$. In the experiments, we will take the EOC to be at the current level and the previous refinement.

All the experiments have been implemented under the Distributed and Unified Numerics Environment (DUNE) [2, 9].

3.6.1 Flat case experiment

The first example is for a flat domain. Let Ω be the unit disc in $\mathbb{R}^2$ centred at the origin and $\mathcal{P} := \{(0,0), (0.5,0), (-0.5,0), (0,0.5), (0,-0.5)\}$ be 5 distinct points in Ω. The PDE boundary value problem is

$$\Delta^2 u = 0 \text{ in } \Omega \setminus \mathcal{P}$$

such that

$$u(X_j) = g(X_j) \;\forall X_j \in \mathcal{P},\; u|_{\partial\Omega} = \Delta u|_{\partial\Omega} = 0,$$

where $g(x) := 1 - |x|^2 + \frac{|x|^2}{2}\log(|x|^2)$. It has the solution

$$u(x) = 1 - |x|^2 + \frac{|x|^2}{2}\log(|x|^2).$$

This can be viewed as a flat biomembrane problem with $\kappa = 1$ and $\sigma = 0$. The coupled second order system is

$$\begin{aligned}
2\Delta u - u - \Delta w + w &= 0 \text{ in } \Omega \setminus \mathcal{P},\\
-\Delta u + u - w &= 0 \text{ in } \Omega,\\
u|_{\partial\Omega} = w|_{\partial\Omega} &= 0,\\
u(X_j) &= g(X_j) \;\forall X_j \in \mathcal{P}.
\end{aligned}$$

As in Subsection 3.3.2, we see that for the first equation, this is not posed on the domain Ω, but away from the points being constrained.

The bilinear forms become

$$c(u,\eta)=\int_\Omega -2\nabla u\cdot\nabla\eta - u\eta,\quad b(u,\eta)=\int_\Omega \nabla u\cdot\nabla\eta+u\eta,\quad m(w,\xi)=\int_\Omega w\xi.$$

Since the problem is posed with homogeneous Navier boundary conditions on the unit disc, we may pose the discrete problem on a polygonal domain Ω_h which approximates the unit disc from within and extend the finite element spaces to be 0 in the skin $\Omega\backslash\Omega_h$. We only calculate the error on the discrete domain, it is clear that the error due to the skin will be sufficiently small that it should not interfere with the calculated interior error. Errors are displayed in the Tables 3.1, 3.2 and 3.3. The errors of $\|w-w_h\|_{0,2}$, $\|w-w_h\|_{1,\frac{4}{3}}$ and $\|u-u_h\|_{1,2}$ behave as expected from the theory provided in Section 3.3, converging at rate almost h, $h^{1/2}$ and h respectively. Whereas the errors $\|u-u_h\|_{0,2}$ and $\|\lambda-\lambda_h\|_{\mathbb{R}^5}$ appear to converge at a higher rate. The theory provided in Section 3.3 suggests that $\|u-u_h\|_{0,2}$ and $\|\lambda-\lambda_h\|_{\mathbb{R}^5}$ would converge at order arbitrarily close to h.

One might hope to use a duality argument to demonstrate the convergence observed, that is to say, let $(\psi,\phi,\chi)\in W^{1,q}(\Omega)\times W^{1,p}(\Omega)\times\mathbb{R}^5$ satisfy

$$\begin{aligned} c(\eta,\psi)+b(\eta,\phi)+(T\eta,\chi)_{\mathbb{R}^N} &= (u-u_h,\eta)_{L^2(\Omega)} \quad \forall \eta\in W^{1,q}(\Omega),\\ b(\psi,\xi)-m(\phi,\xi)&=0 \quad \forall\xi\in W^{1,p}(\Omega),\\ T\psi&=0, \end{aligned}$$

which has the estimate $\|\psi\|_{2,2}+\|\phi\|_{1,p}+\|\chi\|_{\mathbb{R}^5}\leq C\|u-u_h\|_{0,2}$. By testing the dual system with $(u-u_h,w-w_h,\lambda-\lambda_h)$ and summing, one obtains

$$\|u-u_h\|_{0,2}^2=c(u-u_h,\psi)+b(u-u_h,\phi)+b(\psi,w-w_h)-m(\phi,w-w_h),$$

where we have made use of $T(u-u_h)=0$. It is useful to recall that $c(u-u_h,\psi_h)+b(u-u_h,\phi_h)+b(\psi_h,w-w_h)-m(\phi,w-w_h)=0$ for any $(\psi_h,\phi_h)\in\mathcal{S}_h\times\mathcal{S}_h$ with $T_h\psi_h=0$, which leads to

$$\begin{aligned}\|u-u_h\|_{0,2}^2 =& c(u-u_h,\psi-\psi_h)+b(u-u_h,\phi-\phi_h)+b(\psi-\psi_h,w-w_h)-m(\phi-\phi_h,w-w_h)\\ \leq& C\left(\|u-u_h\|_{1,q}+\|w-w_h\|_{1,p}\right)\left(\|\psi-\psi_h\|_{1,q}+\|\phi-\phi_h\|_{1,p}\right).\end{aligned}$$

By choosing ψ_h and ϕ_h to solve the discrete dual equation, one obtains the following estimate

$$\begin{aligned}\|u-u_h\|_{0,2}^2 \leq& C\|u-u_h\|_{0,2}\left(h^{2/q}+|\log(h)|h^{\min(2/q,1-2/q)}\right)\left(h^{2/q}+|\log(h)|h^{\min(2/q,1-2/q)}\right)\\ \leq& C|\log(h)|^2h^{\min(4/q,1,2-4/q)}.\end{aligned}$$

One may see that for $q=4$, the above has an evaluation of $|\log(h)|^2h$. This is estimate is not significantly different to the estimate inherited from the bound on $\|u-u_h\|_{1,2}$.

h	E_{L^2}	E_{H^1}	EOC_{L^2}	EOC_{H^1}
0.420334	0.0347383	0.132332	–	–
0.221925	0.010496	0.0724977	1.87385	0.943152
0.113732	0.00293398	0.0377392	1.90671	0.976601
0.0575358	0.000787479	0.0191858	1.93016	0.992797
0.0289325	0.000206736	0.00965453	1.94547	0.998988

Table 3.1: Errors and experimental orders of convergence for $u - u_h$ in the flat case experiment, Subsection 3.6.1.

h	E_{L^2}	$E_{W^{1,\frac{4}{3}}}$	EOC_{L^2}	$EOC_{W^{1,\frac{4}{3}}}$
0.420334	0.0242845	0.435937	–	–
0.221925	0.0114147	0.316908	1.18197	0.499267
0.113732	0.0057894	0.228215	1.01552	0.491137
0.0575358	0.00292883	0.162944	0.99999	0.494371
0.0289325	0.00147189	0.115812	1.0009	0.496675

Table 3.2: Errors and experimental orders of convergence for $w - w_h$ in the flat case experiment, Subsection 3.6.1.

3.6.2 Surface numerical experiment

The second numerical example is for the surface of the unit sphere, $\Gamma := \mathbb{S}(0,1)$. The point constraints are fixed at the six distinct points $\mathcal{P} := \{(\pm 1,0,0),(0,\pm 1,0),(0,0,\pm 1)\}$. We consider the problem of $\kappa = \sigma = R = 1$ in the forms defined in Definition 3.5.1 corresponding to, is to the PDE boundary value problem, find $(u,\overline{p})$ such that

$$\Delta_\Gamma^2 u + \Delta_\Gamma u - 2u + \overline{p} = f - \Delta_\Gamma g + g \text{ in } \Gamma \setminus \mathcal{P},$$
$$u(X_j) = Z_j \ \forall X_j \in \mathcal{P},$$
$$\int_\Gamma u = 0,$$

where

$$f = 9x_3 \log(1-x_3) + 9x_3 - 2\log(1-x_3) + \frac{1}{2}(5 + 3\log(4)), \quad Z_j = U(X_j) \ \ j = 1,2,...,6,$$

$$g = -3x_3 \log(1-x_3) - 3x_3 - \frac{1}{2}(\log(4)+1)$$

and

$$U = (1-x_3)\log(1-x_3) - \frac{1}{2}(\log(4)-1).$$

We recall that $\overline{p}$ arises as the Lagrange multiplier associated to the constraint $\int_\Gamma u = 0$, as in Subsection 2.1. The solution to this problem is $u = U$, $\overline{p} = 0$. The second order splitting system

h	E_{ℓ^2}	EOC_{ℓ^2}
0.420334	0.00621158	–
0.221925	0.00492358	0.363827
0.113732	0.00218382	1.2161
0.0575358	0.000763246	1.5427
0.0289325	0.000235913	1.70795

Table 3.3: Errors and experimental order of convergence for $\lambda - \lambda_h$ in the flat case experiment, Subsection 3.6.1.

is taken to be

$$\begin{aligned} 3\Delta_\Gamma u - 3u - \Delta_\Gamma w + w + \overline{p} &= f \text{ in } \Gamma \setminus \mathcal{P}, \\ -\Delta_\Gamma u + u - w + \overline{q} &= g \text{ in } \Gamma, \\ u(X_j) &= Z_j \ \forall X_j \in \mathcal{P}, \\ \int_\Gamma u = \int_\Gamma w &= 0, \end{aligned}$$

where $\overline{q}$ is the Lagrange multiplier due to the constraint on the mean value of w.

Thus the forms of Definition 3.3.1 with $\kappa = \sigma = R = 1$ are given by

$$b(u,\eta) = \int_\Gamma \nabla_\Gamma u \cdot \nabla_\Gamma \eta + u\eta, \ c(u,\eta) = -3b(u,\eta), \ m(w,\xi) = \int_\Gamma w\xi.$$

The well-posedness of the problem follows from Section 3.3 and has solution $u = U$, $w = \log(1 - x_3)$.

In these numerical computations, implementation of the point constraints is achieved via penalty with $\epsilon = 10^{-8}$, a value chosen sufficiently small as to play no role in the investigation of the order of convergence with respect to h. The errors are displayed in Tables 3.4 and 3.5. They behave similarly to that of the flat case experiment and are consistent with the theory provided in Section 3.3.

h	E_{L^2}	E_{H^1}	EOC_{L^2}	EOC_{H^1}
0.311152	0.012565	0.0841661	–	–
0.156914	0.00356525	0.042819	1.84007	0.987187
0.0786276	0.000990194	0.0215476	1.85403	0.993838
0.0393352	0.000276744	0.0107968	1.84061	0.997706
0.0196703	7.88541e-05	0.00540193	1.81165	0.999252

Table 3.4: Errors and experimental orders of convergence for $u - u_h^l$ in the surface numerical experiment, Subsection 3.6.2.

h	E_{L^2}	$E_{W^{1,\frac{4}{3}}}$	EOC_{L^2}	$EOC_{W^{1,\frac{4}{3}}}$
0.311152	0.0486308	0.236187	–	–
0.156914	0.0212111	0.165895	1.21203	0.516039
0.0786276	0.0098867	0.118446	1.10472	0.487569
0.0393352	0.00478169	0.0845555	1.04879	0.486638
0.0196703	0.00235552	0.0602071	1.02167	0.49006

Table 3.5: Errors and experimental orders of convergence for $w - w_h^l$ in the surface numerical experiment, Subsection 3.6.2.

3.6.3 Penalty experiment

We now fix h to be sufficiently small that it should contribute little error and take a sequence of ϵ which will tend to 0. For simplicity, we consider the same experiment as in Subsection 3.6.2. Where previously the quantities E and EOC have been functions of h, they will now be functions of ϵ in the expected way. The grid is fixed to be the smallest grid used in the previous experiment with $h = 0.0196703$. In Tables 3.6, 3.7 and 3.8 we see that the errors are consistent with the results of Corollary 3.3.13, Theorem 3.5.4 and Theorem 3.5.8.

ϵ	E_{L^2}	E_{H^1}	EOC_{L^2}	EOC_{H^1}
0.2	0.178146	0.173119	–	–
0.1	0.09091	0.0910002	0.970551	0.970551
0.05	0.0462214	0.0477307	0.975878	0.930951
0.025	0.0233842	0.025107	0.983028	0.926831
0.0125	0.0117769	0.013647	0.989572	0.879506

Table 3.6: Errors and experimental orders of convergence for $u - (u_h^\epsilon)^l$ in the numerical experiment, Subsection 3.6.3.

ϵ	E_{L^2}	$E_{W^{1,\frac{4}{3}}}$	EOC_{L^2}	$EOC_{W^{1,\frac{4}{3}}}$
0.2	0.337878	0.441999	–	–
0.1	0.191338	0.29658	0.820381	0.575622
0.05	0.107691	0.19941	0.829231	0.572684
0.025	0.0592661	0.134285	0.861611	0.570441
0.0125	0.0316392	0.0952299	0.905497	0.495808

Table 3.7: Errors and experimental orders of convergence for $w - (w_h^\epsilon)^l$ in the numerical experiment, Subsection 3.6.3.

3.6.4 Surface numerical and penalty experiment

We now couple ϵ and h, we take $\epsilon \approx Ch^2$. The same experiment as in Subsections 3.6.2 and 3.6.3 is used. The E and EOC are calculated in terms of the grid size h. In Tables 3.9 and 3.10

ϵ	E_{ℓ^2}	EOC_{ℓ^2}
0.2	0.568092	–
0.1	0.421624	0.430169
0.05	0.287481	0.552491
0.025	0.177819	0.693052
0.0125	0.101217	0.81296

Table 3.8: Errors and experimental order of convergence for $\lambda - \frac{T_h u_h^\epsilon - Tu}{\epsilon}$ in the numerical experiment, Subsection 3.6.3.

we see that the errors are consistent with the results of of Corollary 3.3.13, Theorem 3.5.4 and Theorem 3.5.8.

h	ϵ	E_{L^2}	E_{H^1}	EOC_{L^2}	EOC_{H^1}
0.311152	0.2	0.182674	0.193639	–	–
0.156914	0.05	0.0471203	0.063719	1.97931	1.62364
0.0786276	0.0125	0.0118776	0.0248031	1.99435	1.36548
0.0393352	0.003125	0.00296094	0.0112296	2.00569	1.14411
0.0196703	0.00078125	0.000776274	0.00546128	1.9318	1.0402

Table 3.9: Errors and experimental orders of convergence for $u - (u_h^\epsilon)^l$ in the numerical experiment, Subsection 3.6.4.

h	ϵ	E_{L^2}	E_{H^1}	EOC_{L^2}	$EOC_{W^{1,\frac{4}{3}}}$
0.311152	0.2	0.337009	0.488743	–	–
0.156914	0.05	0.104921	0.253073	1.70455	0.961402
0.0786276	0.0125	0.0310131	0.140968	1.76387	0.846845
0.0393352	0.003125	0.00893405	0.0271549	1.79691	0.667151
0.0196703	0.00078125	0.00300472	0.00677873	1.57239	0.54459

Table 3.10: Errors and experimental orders of convergence for $w - (w_h^\epsilon)^l$ in the numerical experiment, Subsection 3.6.4.

3.7 Conclusion

In this chapter, we have demonstrated the well-posedness of a generalised saddle point problem with either linear constraints or penalising linear constraints. With this we have given an abstract numerical analysis. We have shown how we may apply this abstract theory to a flat biomembrane problem and a near-spherical biomembrane problem with point constraints. We have given numerical experiments, demonstrating the numerical convergence in terms of the grid size and penalty parameter.

Chapter 4

Differentiability of the membrane energy

4.1 Introduction

We recall that the lipid bilayer which forms a biomembrane is believed to act like a fluid in the lateral direction and elastically in the normal direction. This means that in principle, any proteins which may be embedded into or attached to the surface of the membrane may move freely. This means that not only can the proteins influence the shape of the membrane, but also the protein interaction will be membrane mediated.

Indeed, although direct protein-protein interactions are important, [48] demonstrated that the long range interactions are predominantly membrane mediated. An overview of membrane mediated interactions is given in [8]. An assumption of symmetry of the protein inclusion allows for either analytic representation or approximation by an asymptotic expansion of the interactions [59, 79, 27, 81, 44]. Frequently the studies of these interactions were restricted to a nearly flat membrane with circular or single point inclusions. It is known that the shape of the inclusion has a significant impact on the interaction [60]. In the recent work of [76], they consider a near spherical membrane which is deformed by particles which attach along segments of an ellipsoid or hyperbolid and in [50], they consider arbitrary, sufficiently regular, particle inclusions on a flat membrane. Recent work has looked at shape formation of multiple smaller particles into larger structures [78, 49]. The article [25] considers generic elastic energies on a manifold with embedded point particles which have a given interaction potential. A variational formulation for equilibria of the surface and particle system is presented, along a discretisation. Numerical validations are given, in particular, a Helfrich problem is presented. We further note the work of [14] which considers point constraints in a Kirchoff plate. This bears a striking similarity to the biological problems of optimising the locations of constraints with respect to the an elastic membrane energy.

We assume that the attached proteins are rigid, that is to say they do not bend and can only move by translations or rotations. It is of clear interest to consider the force that the membrane exerts on these attached proteins. This is relevant to, say, calculate locally minimising configuration of multiple proteins via a gradient flow, to estimate statistical quantities using over-damped Langevin Dynamics [66, Section 2.2.2] or as a step for a full model for the problem of particles in membranes. For further details on estimation of the free energy of a particle membrane, see [58].

The derivative of the energy with respect to particle location is calculated as a shape derivative in [34], and appears by use of a pull back method in [50], both in the case of large particles on a nearly flat membrane. We will follow many of the ideas of this second work, making use of methods from [18] to deal with the fact we are on a surface rather than a flat domain.

One motivation for constructing a formula for the membrane mediated particle interactions may be seen from the following example. For $\overline{\mathcal{E}}(p)$ the total energy of the particle system (the membrane energy with electrostatic interaction) in configuration p, one might be interested in finding p^* such that $\overline{\mathcal{E}}(p^*)$ is minimal. One may choose to do this with a gradient descent algorithm in which an update step might be:

$$p_{n+1} = p_n - \alpha_n \nabla_p \overline{\mathcal{E}}(p_n),$$

for some $\alpha_n > 0$ which may depend on n. Clearly one may approximate the derivative $\nabla_p \overline{\mathcal{E}}(p_n)$ by taking a difference quotient. However this will be expensive, as one would require solving $3N + 1$ linear systems - the system associated to the state p_n and the $3N$ directions that ∇_p corresponds to. With the explicit formula we find, the algorithm to construct the gradient would require solving 1 linear system and evaluating $3N$ functionals, where these functionals relatively cheap to evaluate compared to a linear solve for a fourth order PDE.

4.1.1 Outline

The formula for the derivative of the minimising energy with respect to the location of the particles is derived in Section 4.2. Some numerical examples are presented in Section 4.3. In a finite element setting we calculate and compare derivatives using the formula and a difference quotient of the energies for comparison.

4.2 Gradient of the energy with respect to configuration changes

In this section we find a formula for the derivative of $\mathcal{E}(p)$ with with respect to changes in the configuration p.

Definition 4.2.1 (Derivative of the configurational energy)**.** *The configurational energy is dif-*

ferentiable at $p \in \Lambda^\circ$ in the direction $e \in \prod_{i=1}^N (\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma)$ if the derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(p+te)|_{t=0},$$

exists and we denote this by $\partial_e \mathcal{E}(p)$.

The difficulty lies in the implicit definition of the energy $\mathcal{E}(p)$ in terms of the minimisation of the quadratic energy $J(v)$ over the configurational space $U(p)$ requiring the evaluation of

$$\frac{\mathrm{d}}{\mathrm{d}t}J(u(p+te))|_{t=0}$$

which involves the minimisation of $J(\cdot)$ over $U(p+te)$. In order to achieve this we fix p and employ suitable local isomorphisms on the vector spaces $U(p)$ via appropriate diffeomorphisms of the domain $\Gamma(p)$. This is applied locally to transform the energy (2.1) and the related minimisation problems over a reference function space.

We make the following assumption:

Assumption 4.2.2. *Let $k \geq 1$. For each $p \in \Lambda^\circ$ there exists an open ball $\mathcal{B} \subset \prod_{i=1}^N \left(\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma\right)$ containing 0 and a family of C^k-diffeomorphisms $\chi \colon \mathcal{B} \times \Gamma \to \Gamma$ such that*

$$\chi(0,\cdot) \text{ is the identity on } \Gamma$$

and for all $q \in \mathcal{B}$, $p + q \in \Lambda^\circ$ and

$$v \circ \chi(q,\cdot)^{-1} \in U(p+q) \iff v \in U(p). \tag{4.1}$$

This assumption is verified by construction in Appendix D. We now define what we mean by the derivative of χ with respect to e.

Definition 4.2.3. *Given $q \in \mathcal{B}$ and $e \in \prod_{i=1}^N \left(\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma\right)$, for each $x \in \Gamma$, the derivative of $\chi(\cdot, x)$ at q in direction e is defined to be*

$$\partial_e \chi(q,x) := \frac{\mathrm{d}}{\mathrm{d}t}\chi(q+te,x)|_{t=0}.$$

Remark 4.2.4. *Notice that:*

- *The dependence on p of $\mathcal{B}$ and χ has been suppressed.*
- *For our purposes we will not require full knowledge of the diffeomorphism χ, only the derivative $\partial_e \chi(0,\cdot)$.*
- *The fact that Λ may be identified as a subset of the finite dimensional space $\mathbb{R}^{3\times N}$ will be exploited to reduce the problem of differentiability of $\mathcal{E}$ to be an application of the Implicit Function Theorem applied to a reformulated interaction energy.*

- *The condition* (4.1) *may be decomposed into three parts:* $T(p+q)\left(v\circ\chi^{-1}\right) = T(p)v$ *for all* $v \in H^2(\Gamma)$, $\int_\Gamma v\circ\chi^{-1} = \int_\Gamma v$ *for all* $v\in H^2(\Gamma)$ *and* $v \in H^2(\Gamma) \iff v\circ\chi^{-1}\in H^2(\Gamma)$.
- *The condition on* χ *that* $\int_\Gamma v\circ\chi(q,\cdot)^{-1} = \int_\Gamma v$ *for all* $v \in H^2(\Gamma)$ *is equivalent to requiring that* $\det\left(\nabla_\Gamma\chi(q,\cdot)+\nu(\cdot)\circ\chi(q,\cdot)\otimes\nu(\cdot)\right) = 1$ *on* Γ. *As such, it is sufficient to have* $\partial_e \det\left(\nabla_\Gamma\chi(q,\cdot)+\nu(\cdot)\circ\chi(q,\cdot)\otimes\nu(\cdot)\right) = 0$ *for any* $e\in\prod_{i=1}^N\left(\mathbb{R}\times T_{X_{\mathcal{G}_i}}\Gamma\right)$. *We will later see that, for* $q=0$, *this is the same as requiring* $\mathrm{div}_\Gamma\partial_e\chi(0,\cdot)$ *vanishes.*

4.2.1 The transformed functional and its derivative

Using the χ satisfying Assumption 4.2.2, we have the following functional.

Definition 4.2.5. *Let* $J^*\colon \mathcal{B}\times U(p)\to\mathbb{R}$ *be given by*

$$J^*\colon (q,v)\mapsto J(v\circ\chi^{-1}(q,\cdot)),\ J^*(0,v) = J(v)\ \forall v\in U(p).$$

We call $J^*(\cdot,v)$ *the transformed membrane energy. Given* $e\in\prod_{i=1}^N\left(\mathbb{R}\times T_{X_{\mathcal{G}_i}}\right)$, *if, for any* $v\in U(p)$, *the derivative*

$$\frac{\mathrm{d}}{\mathrm{d}t}J^*(te,v)|_{t=0}$$

exists, we denote it $\partial_e J^*(0,v)$.

We now define some terms which appear in [18] which are useful to give an explicit representation of J^*.

Definition 4.2.6. *Given* $q\in\mathcal{B}$, *we define on* Γ *the matrices and determinant*

$$\begin{aligned} B =& B(q,\cdot) := \nabla_\Gamma\chi(q,\cdot)+\nu(\cdot)\circ\chi(q,\cdot)\otimes\nu(\cdot),\\ G =& G(q,\cdot) := B(q,\cdot)^T B(q,\cdot),\\ b =& b(q) := \det(B(q,\cdot)). \end{aligned}$$

The following, convenient representation of J^* is immediate from the results Lemmas A.0.1 and A.0.2 in the appendix.

Lemma 4.2.7. *Given* $v\in U(p)$, $q\in\mathcal{B}$, *it holds that*

$$J^*(q,v) = \frac{\kappa}{2}\int_\Gamma \frac{1}{b}\left(\mathrm{div}_\Gamma(bG^{-1}\nabla_\Gamma v)\right)^2 + \left(\frac{\sigma}{2}-\frac{\kappa}{R^2}\right)\int_\Gamma b\nabla_\Gamma v\cdot G^{-1}\nabla_\Gamma v - \frac{\sigma}{R^2}\int_\Gamma bv^2. \tag{4.2}$$

Note that we wish to differentiate J^* with respect to q and that the q dependence is located in the coefficients $B(q)$.

Lemma 4.2.8. *Suppose* $\mathcal{B}\subset\prod_{i=1}^N(\mathbb{R}\times T_{X_{\mathcal{G}_i}}\Gamma)$ *is sufficiently small with* $0\in\mathcal{B}$ *and* $\chi\in C^k(\mathcal{B}\times\Gamma;\Gamma)$, *then* $J^*\in C^{k-2}(\mathcal{B}\times U(p);\mathbb{R})$.

Proof. It is clear from the expression for J^* that it depends on B, the derivative of B and smoothly (in $H^2(\Gamma)$) on v. Since $B(0) = \mathrm{I}$, the identity matrix, B depends continuously on q and det is a continuous map, thus for a sufficiently small neighbourhood $\mathcal{B} \ni 0$, $\det(B(q)) > c > 0$ it hold that B is non-singular. Thus by smoothness of the integrand, we may apply the dominated convergence theorem to obtain $J^* \in C^{k-2}(\mathcal{B} \times U(p); \mathbb{R})$. $\square$

Theorem 4.2.9. *There exists an open neighbourhood $\hat{\mathcal{B}}$ of 0 in $\prod_{i=1}^N (\mathbb{R} \times T_{X_{\mathcal{G}_i}} \Gamma)$ such that $\mathcal{E}(p + \cdot) \in C^{k-2}(\hat{\mathcal{B}}; \mathbb{R})$. In particular, for $k \geq 3$ and $u = \operatorname{argmin}_{v \in U(p)} J(v)$,*

$$\partial_e \mathcal{E}(p) = \partial_e J^*(0, u).$$

Proof. In the following we suppress the dependence on p and write $u = u(p), U_0 = U_0(p)$. Define $\mathcal{J} \in C^{k-2}(\mathcal{B} \times U_0; \mathbb{R})$ by

$$\mathcal{J}(q, v) := J^*(q, u + v) \quad \text{for} \quad (q, v) \in \mathcal{B} \times U_0.$$

For fixed q, $\mathcal{J}(q, \cdot)$ is a quadratic functional and by the definition of u we have that the minimiser of the functional $\mathcal{J}(0, v)$ over U_0 is given by $v = 0$. Define $F \in C^{k-2}(\mathcal{B} \times U_0; U_0^*)$ by

$$F(q, v) := D_v \mathcal{J}(q, v)$$

where, for fixed q, $D_v \mathcal{J}$ is the first variation of $\mathcal{J}(q, \cdot)$ over U_0. For each (q, v), $F(q, v)$ is a linear functional. Since $J(0, v)$ attains minima at $v = 0$, it follows that $F(0, 0) = D_v \mathcal{J}(0, 0) = 0 \in U_0^*$. Furthermore, the first variation of F at $(0, 0)$,

$$D_v F(0, 0) \colon (\xi, \eta) \in U_0 \times U_0 \mapsto D_v F(0, 0)[\xi, \eta] = D_{vv} \mathcal{J}(0, 0)[\xi, \eta] = a(\xi, \eta),$$

is a strictly coercive bilinear form over $U_0 \times U_0$. As a consequence, it follows that the map $U_0 \ni v \mapsto D_v F(0, v) \in U_0^*$ is invertible.

It therefore holds that we may apply the implicit function theorem, Theorem B.0.1, to $f = F$, with $(a, b) = (0, 0)$, $\mathcal{X} = \prod_{i=1}^N (\mathbb{R} \times T_{X_{\mathcal{G}_i}} \Gamma)$, $\mathcal{Y} = U_0$, $\mathcal{Z} = \mathcal{Y}^*$ and $\Omega = \mathcal{B} \times \mathcal{Y}$. As such, there is neighbourhood of 0, $\hat{\mathcal{B}} = V \subset \mathcal{B}$ and a function $\hat{v} \in C^{k-2}(\hat{\mathcal{B}}; U_0(p))$ such that $\hat{v}(0) = 0$ and $F(q, \hat{v}(q)) = 0$. That is to say $J_v^*(q, \hat{v}(q) + u) = 0$, so $\hat{v}(q) + u$ is a critical point of $J^*(q, \cdot)$. By coercivity of $J^*(q, \cdot)$ over $U(p)$, $\hat{u}(q) := \hat{v}(q) + u$ is the unique minimiser. Hence

$$\mathcal{E}(p + q) = \min_{\eta \in U(p+q)} J(\eta) = \min_{\eta \in U(p)} J^*(q, \eta) = J^*(q, \hat{u}(q)).$$

Since $\hat{u} \in C^{k-2}(\hat{\mathcal{B}}; U(p))$, $J^* \in C^{k-2}(\mathcal{B} \times U(p); \mathbb{R})$, it follows $\mathcal{E}(p + \cdot) \in C^{k-2}(\hat{\mathcal{B}}; \mathbb{R})$. Taking the derivative of $\mathcal{E}$ gives

$$\partial_e \mathcal{E}(p) = \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(p + te)|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} J^*(te, u)|_{t=0} + \frac{\mathrm{d}}{\mathrm{d}t} J^*(0, \hat{u}(te))|_{t=0} = \partial_e J^*(0, u),$$

where $\frac{\mathrm{d}}{\mathrm{d}t}J^*(0,\hat{u}(te))|_{t=0} = D_v J^*(0,u)\left[\frac{\mathrm{d}}{\mathrm{d}t}\hat{u}(te)|_{t=0}\right]$ vanishes since $D_v J^*(0,u) = 0$. $\square$

Remark 4.2.10. *Although J^* depends on the choice of χ, the derivative $\partial_e \mathcal{E}(p)$ is independent of the choice of χ. One may consider a different diffeomorphism, say, $\tilde{\chi}$ with energy $\tilde{J}^*$, one would then have that*

$$\min_{\eta \in U(p+q)} J^*(q,\eta) = \min_{\tilde{\eta} \in U(p+q)} \tilde{J}^*(q,\tilde{\eta})$$

and arrive at $\partial_e \mathcal{E}(p) = \partial_e \tilde{J}^*(0,u) = \partial_e J^*(0,u)$.

4.2.2 An explicit formula for the derivative

It is convenient to define the following.

Definition 4.2.11. *Define the tangential vector field $V \colon \prod_{i=1}^N \left(\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma\right) \times \Gamma \to \mathbb{R}^3$ by*

$$V(e,x) := \partial_e \chi(0,x),$$

which is tangential in the sense that $V(e,x) \in T_x\Gamma$ for all $(e,x) \in \prod_{i=1}^N \left(\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma\right) \times \Gamma$.

Proposition 4.2.12. *Given $e \in \prod_{i=1}^N \left(\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma\right)$, set $\mathcal{A} := (\mathrm{div}_\Gamma V)I - (\nabla_\Gamma V + \nabla_\Gamma V^T)$ then for $\eta \in H^2(\Gamma)$*

$$\begin{aligned}\partial_e J^*(0,\eta) =& \kappa \int_\Gamma (\mathcal{A} : D^2_\Gamma \eta - \Delta_\Gamma V \cdot \nabla_\Gamma \eta)\Delta_\Gamma \eta \\ &- \frac{\kappa}{R^2}\int_\Gamma (V \cdot \nabla_\Gamma \eta + \frac{1}{2}\mathrm{div}_\Gamma V \Delta_\Gamma \eta)\Delta_\Gamma \eta \\ &+ \left(\frac{\sigma}{2} - \frac{\kappa}{R^2}\right)\int_\Gamma \nabla_\Gamma \eta \cdot \mathcal{A}\nabla_\Gamma \eta - \frac{\sigma}{R^2}\int_\Gamma \mathrm{div}_\Gamma V \eta^2.\end{aligned}$$

Proof. We will make use of the fact that $B(0) = \mathrm{I}$ and $\det(B(0)) = 1$. To simplify notation when taking derivative ∂_e, we assume that we are evaluating at $q = 0$, if there is no argument given. The product rule gives

$$\begin{aligned}\partial_e J^*(0,\eta) =& \frac{\kappa}{2}\int_\Gamma 2\mathrm{div}_\Gamma \frac{\mathrm{d}}{\mathrm{d}t}(\det(B(te))G(te)^{-1}\nabla_\Gamma \eta)|_{t=0}\Delta_\Gamma \eta - (\Delta_\Gamma \eta)^2 \frac{\mathrm{d}}{\mathrm{d}t}\det(B(te))|_{t=0} \\ &+ \left(\frac{\sigma}{2} - \frac{\kappa}{R^2}\right)\int_\Gamma \nabla_\Gamma \eta \cdot \frac{\mathrm{d}}{\mathrm{d}t}(\det(B(te))G(te)^{-1})|_{t=0}\nabla_\Gamma \eta \\ &- \frac{\sigma}{R^2}\int_\Gamma \frac{\mathrm{d}}{\mathrm{d}t}\det(B(te))|_{t=0}\eta^2.\end{aligned} \tag{4.3}$$

Where we calculate

$$\begin{aligned}\partial_e B =& \nabla_\Gamma V + (\mathcal{H}V)\otimes\nu,\\ \partial_e \det(B) =& \mathrm{div}_\Gamma V,\\ \partial_e B^{-1} =& -\nabla_\Gamma V - (\mathcal{H}V)\otimes\nu.\end{aligned}$$

Since $G := B^T B$ one has,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\det(B(te))G(te)^{-1})|_{t=0} = (\mathrm{div}_\Gamma V)\,\mathrm{I} - \nabla_\Gamma V - (\mathcal{H}V)\otimes\nu - \nabla_\Gamma V^T - \nu\otimes(\mathcal{H}V).$$

We are also required to calculate the surface divergence of the above quantity

$$\begin{aligned}&\mathrm{div}_\Gamma \partial_e(\det(B)G^{-1})\\ &\qquad =\mathrm{div}_\Gamma\big((\mathrm{div}_\Gamma V)\,\mathrm{I} - \nabla_\Gamma V - (\mathcal{H}V)\otimes\nu - \nabla_\Gamma V^T - \nu\otimes(\mathcal{H}V)\big)\\ &\qquad =\sum_{k=1}^{n+1}(\nabla_\Gamma \underline{D}_k - \underline{D}_k\nabla_\Gamma)V_k - \Delta_\Gamma V - \mathrm{div}_\Gamma((\mathcal{H}V)\otimes\nu + \nu\otimes(\mathcal{H}V))\\ &\qquad =-\Delta_\Gamma V - H\mathcal{H}V + \sum_{k=1}^{n+1}(\mathcal{H}\nabla_\Gamma V_k)_k\nu - (\mathcal{H}\nabla_\Gamma V_k)\nu_k - \underline{D}_k(\nu(\mathcal{H}V)_k).\end{aligned}$$

It is possible to calculate

$$\sum_{k=1}^{n+1}(\mathcal{H}\nabla_\Gamma V_k)_k = \sum_{k=1}^{n+1}\sum_{l=1}^{n+1}\mathcal{H}_{lk}\underline{D}_l V_k = \mathcal{H}:\nabla_\Gamma V,$$

by using that $V\cdot\nu = 0$,

$$\sum_{k=1}^{n+1}(\mathcal{H}\nabla_\Gamma V_k)_j\nu_k = \sum_{k=1}^{n+1}\sum_{l=1}^{n+1}\mathcal{H}_{lj}\underline{D}_l V_k\nu_k = \sum_{k=1}^{n+1}\sum_{k=1}^{n+1}\mathcal{H}_{lj}\underline{D}_l(V_k\nu_k) - \mathcal{H}_{lj}\mathcal{H}_{lk}V_k = -(\mathcal{H}^2V)_j.$$

Furthermore,

$$\begin{aligned}\sum_{k=1}^{n+1}\underline{D}_k(\nu(\mathcal{H}V)_k) =& \sum_{k=1}^{n+1}\sum_{l=1}^{n+1}\underline{D}_k(\nu\mathcal{H}_{kl}V_l) = \sum_{k=1}^{n+1}\sum_{l=1}^{n+1}\underline{D}_k\nu\mathcal{H}_{kl}V_l + \nu\underline{D}_k(\mathcal{H}_{kl}V_l)\\ =& \sum_{k=1}^{n+1}\sum_{l=1}^{n+1}\underline{D}_k\nu\mathcal{H}_{kl}V_l + \nu\underline{D}_k\mathcal{H}_{kl}V_l + \nu\mathcal{H}_{kl}\underline{D}_k V_l\\ =& \mathcal{H}^2V + (\mathcal{H}:\nabla_\Gamma V + (\nabla_\Gamma\cdot\mathcal{H})\cdot V)\nu.\end{aligned}$$

Together this gives,

$$\mathrm{div}_\Gamma(\partial_e(\det(B)G^{-1})) = -\Delta_\Gamma V - \nu(\nabla_\Gamma\cdot\mathcal{H})\cdot V - H\mathcal{H}V,$$

where the middle term will vanish when multiplied against a tangential vector field. We are left with

$$\begin{aligned}\partial_e(\det(B(q))G^{-1}) : D^2_\Gamma\eta = \mathcal{A} : D^2_\Gamma\eta - (\mathcal{H}V)\otimes\nu : D^2_\Gamma\eta \\ - \nu\otimes(\mathcal{H}V) : D^2_\Gamma\eta,\end{aligned}$$

where one may recall that for b, c vectors and matrix A, $A : (b\otimes c) = b^T Ac$. Thus

$$\partial_e(\det(B)G^{-1}) : D^2_\Gamma\eta = \mathcal{A} : D^2_\Gamma\eta + \mathcal{H}^2\nabla_\Gamma\eta\cdot V,$$

which completes the result when evaluating H and $\mathcal{H}$ for a sphere. □

By Theorem 4.2.9, when evaluating this at the energy minimising membrane, we will obtain the derivative we seek. We notice that it might be convenient to integrate by parts to remove the surface Hessian. This will give an alternate formula which is better suited for the numerical methods considered in [33, 38].

Corollary 4.2.13. *Under the assumptions of Proposition 4.2.12 it may be seen that, for $\eta \in W^{3,p}(\Gamma)$, $p < 2$,*

$$\begin{aligned}\partial_e J^*(q,\eta)|_{q=0} = &-\kappa\int_\Gamma \frac{1}{2}(\mathrm{div}_\Gamma V)(\Delta_\Gamma\eta)^2 + \nabla_\Gamma\Delta_\Gamma\eta\cdot\mathcal{A}\nabla_\Gamma\eta \\ &+\frac{1}{2}\left(\sigma - \frac{2\kappa}{R^2}\right)\int_\Gamma \nabla_\Gamma\eta\cdot\mathcal{A}\nabla_\Gamma\eta - \frac{\sigma}{R^2}\int_\Gamma(\mathrm{div}_\Gamma V)\,\eta^2.\end{aligned} \tag{4.4}$$

Proof. This follows from integration by parts in (4.3) and following through with the proof above. The integration by parts is admissible by the regularity of η. □

By the additional regularity shown in Proposition 2.1.7, we see that we may pick $\eta = \mathrm{argmin}_{v\in U(p)} J(v)$ in the above. This gives the main result of the work which follows from the previous results.

Theorem 4.2.14. *Let $p \in \Lambda^\circ$, $u = \mathrm{argmin}_{v\in U(p)} J(v)$ and $\mathcal{A} := (\mathrm{div}_\Gamma V)I - \nabla_\Gamma V - \nabla_\Gamma V^T$, then*

$$\begin{aligned}\partial_e\mathcal{E}(p) = &-\kappa\int_\Gamma \frac{1}{2}(\mathrm{div}_\Gamma V)(\Delta_\Gamma u)^2 + \nabla_\Gamma\Delta_\Gamma u\cdot\mathcal{A}\nabla_\Gamma u \\ &+\frac{1}{2}\left(\sigma - \frac{2\kappa}{R^2}\right)\int_\Gamma \nabla_\Gamma u\cdot\mathcal{A}\nabla_\Gamma u - \frac{\sigma}{R^2}\int_\Gamma(\mathrm{div}_\Gamma V)\,u^2.\end{aligned} \tag{4.5}$$

Proof. This is an application of Theorem 4.2.9 and Corollary 4.2.13. □

Corollary 4.2.15. *Let $N = 1$, then $\partial_e\mathcal{E}(p) = 0$ for all $p \in \Lambda^\circ$ and directions $e \in \mathbb{R}\times T_{X_\mathcal{G}}\Gamma$.*

Proof. This result follows from the symmetry of the sphere and invariance of J under rotations and translations. □

4.3 Numerical experiments

We are now equipped to present some simulations, but first we discuss the approximation errors which arise in numerical simulations.

Proposition 4.3.1. *Let* $\tilde{u} \in W^{1,\infty}(\Gamma)$ *with* $-\Delta_\Gamma \tilde{u} \in W^{1,2-\delta}(\Gamma)$ *for any* $\delta > 0$. *Then for any* $\epsilon \in (0,1)$, $p \in (1,2)$ *and* $q = p^*$, *there is* $C > 0$ *such that*

$$\begin{aligned}|\partial_e J^*(0,\tilde{u}) - \partial_e \mathcal{E}(p)| \leq C \|\nabla_\Gamma V\|_{0,\infty} \Big(& \|\Delta_\Gamma(u-\tilde{u})\|_{1,p} \|\nabla_\Gamma u\|_{1,q} \\ & + \|\Delta_\Gamma(u-\tilde{u})\|_{0,2} \left(\|\Delta_\Gamma \tilde{u}\|_{0,2} + \|\Delta_\Gamma u\|_{0,2}\right) \\ & + \|\nabla_\Gamma(u-\tilde{u})\|_{0,\frac{2-\epsilon}{1-\epsilon}} \|\Delta_\Gamma \tilde{u}\|_{1,2-\epsilon} \\ & + \|\nabla_\Gamma(u-\tilde{u})\|_{1,2} \left(\|\nabla_\Gamma u\|_{1,2} + \|\nabla_\Gamma \tilde{u})\|_{1,2}\right) \Big).\end{aligned}$$

Proof. This follows from the form $\partial_e J^*$ takes in (4.5) and making use of Hölder inequalities. □

The particular form for the estimate above is chosen so that one may apply the error estimates of [38] making use of a split formulation to approximate u and $-\Delta_\Gamma u + u$ with linear finite elements. There may be different estimates one wishes to show which relate to the formula of Proposition 4.2.12, for example, if one were to use a higher order discretisation of the membrane problem such as the method of [63] which deals with a biharmonic problem on surfaces.

4.3.1 Experiments

We now conduct a selection of numerical experiments. These illustrate the formula and that a method of difference quotients may be unreliable. It is clear that the difference quotient will be slower - one would have to solve (at least) two algebraic systems, whereas when using the formula, a single algebraic system is solved and a functional evaluated.

For all of the experiments we fix $\kappa = \sigma = R = 1$. For the optimal membrane shape, $u(p)$, we approximate it by solving a penalised finite element problem, we call this $u_h(p)$. The penalisation weakly enforces the point constraints and is done in order to ease the linear algebra. We solve a split system for this fourth order problem, the well-posedness and analysis of the system is given in Chapter 3 where the error due to using a penalty formulation is shown to be well controlled. All the experiments have been implemented under the Distributed and Unified Numerics Environment (DUNE) [2, 9].

We begin with an experiment to demonstrate the convergence of the numerical calculation of the formula. This is done by fixing a particle configuration and refining the computational mesh. This experiment is then followed by some experiments where we fix the grid and vary the configuration to verify that the derivative we calculate matches the a difference quotient of the energy. In these experiments we also see that the formula is a better method than using difference quotients.

We now define the quantities which we will calculate in the numerical experiments.

Definition 4.3.2. *Let Γ_h be a connected, polygonal surface approximating Γ and $\mathcal{S}_h$ be the space of linear finite element functions on Γ_h. Given $v_h \in \mathcal{S}_h$ a finite element function, let $w_h \in \mathcal{S}_h$ satisfy*

$$\int_{\Gamma_h} \nabla_{\Gamma_h} v_h \cdot \nabla_{\Gamma_h} \eta_h + v_h \eta_h = \int_{\Gamma_h} w_h \eta_h$$

for all $\eta_h \in \mathcal{S}_h$. We define

$$J_h(v_h) := \frac{1}{2} \int_{\Gamma_h} \kappa (w_h - v_h)^2 + \left(\sigma - \frac{2\kappa}{R^2}\right) |\nabla_{\Gamma_h} v_h|^2 - \frac{2\sigma}{R^2} v_h^2,$$

the discrete analogue of (2.1)*. Define*

$$\mathcal{E}_h(p) := J_h(u_h(p)),$$

the discrete analogue of Definition 2.1.17, where $u_h(p)$ is the minimiser of J_h over $\mathcal{S}_h$ such that $\int_{\Gamma_h} u_h(p) = 0$ and $T(p)(u_h^l(p)) = Z$.

Let $V_h = I_h V$, where V is as in Definition 4.2.11 and $I_h \colon C(\Gamma) \to \mathcal{S}_h$ is the interpolation map. Then define $\mathcal{A}_h := I(\mathrm{div}_{\Gamma_h} V_h) - \nabla_{\Gamma_h} V_h - \nabla_{\Gamma_h} V_h^T$ and

$$\begin{aligned}(\partial_e J^*)_h(v_h) := &-\kappa \int_{\Gamma_h} \frac{1}{2} \left(\mathrm{div}_{\Gamma_h} V\right) (v_h - w_h)^2 + \nabla_{\Gamma_h}(v_h - w_h) \cdot \mathcal{A}_h \nabla_{\Gamma_h} v_h \\ &+ \frac{1}{2}\left(\sigma - \frac{2\kappa}{R^2}\right) \int_{\Gamma_h} \nabla_{\Gamma_h} v_h \cdot \mathcal{A}_h \nabla_{\Gamma_h} v_h - \frac{\sigma}{R^2} \int_{\Gamma_h} \left(\mathrm{div}_{\Gamma_h} V\right) v_h^2,\end{aligned}$$

the discrete analogue of (4.4)*.*

Note that $(\partial_e J^*)_h$ is not necessarily the derivative of $\mathcal{E}_h$. It is clear that the difference quotients we calculate will be approximations of the derivative of $\mathcal{E}_h$, should it exist, but not necessarily close to $(\partial_e J^*)_h$.

For the first three experiments we use $V(\cdot,\cdot) = \mathcal{V}(0,\cdot,\cdot)$ as in the construction in Definition D.0.2. We take δ to be roughly h so that the interpolation of V has support on a small, fixed number of vertices. This makes the evaluation of the functional very quick. For the remaining experiments, V is constructed as in Section D.0.0.1, where the r and ϵ we use for the cut off function are taken to be $r = 0.75$ and $\epsilon = 0.15$.

For the presented convergence experiment, we do not know the exact values of the quantities we estimate. We take the error at level h to be given by the difference between the value at level h and the value on the most refined grid. That is for quantity F_h and smallest grid size h^*, we say the error E_h is given by $|F_h - F_{h^*}|$. For two grids with size h_1 and h_2, we say the EOC of F_h is given by $\log(E_{h_1}/E_{h_2})/\log(h_1/h_2)$, we will take h_1 and h_2 to be from successively refined grids.

h	δ_h	$\mathcal{E}_h(-\theta(\delta_h))$	$\mathcal{E}_h(0)$	$\mathcal{E}_h(\theta(\delta_h))$	$(\partial_e J^*)_h(u_h)$	DQ_h
0.301511	0.25	16.7958	17.199	16.3577	-1.2195	-1.5438
0.152499	0.125	15.524	15.5781	15.3318	-1.33257	-1.4439
0.0764719	0.0625	15.0356	15.0309	14.945	-1.37356	-1.40516
0.0382639	0.03125	14.8615	14.8509	14.8174	-1.38244	-1.39168
0.0191355	0.0078125	14.8006	14.7929	14.7788	-1.38464	-1.3872

Table 4.1: Calculated quantities for experiment in Subsection 4.3.1.1

4.3.1.1 Convergence experiment

We begin by checking the formula and the finite element approximation. We consider 6 particles each consisting of a single point. The points and constraints are given by

$$\begin{aligned} X_1 =&(0,0,1)^T, \; Z_1 = 1; \qquad & X_2 =&(0,0,-1)^T, \, Z_2 = 0; \\ X_3 =&(0,1,0)^T, \; Z_3 = 0; \qquad & X_4 =&(0,-1,0)^T, \, Z_4 = 0; \\ X_5 =&(1,0,0)^T, \; Z_5 = 0.1; \qquad & X_6 =&(-1,0,0)^T, \, Z_6 = 0. \end{aligned}$$

Approximate evaluations of the derivative in the direction $e = (1,0,0)^T \in T_{X_1}\Gamma$ are computed together with approximations of the energy. For each finite element mesh size h, we calculate

$$\mathcal{E}_h(0),\, \mathcal{E}_h(\theta(\delta_h)),\, \mathcal{E}_h(-\theta(\delta_h)),\, (\partial_e J^*)_h(u_h).$$

Here $\mathcal{E}_h(\theta(\delta))$ denotes the energy where the point X_1 is replaced by the point

$$X_1(\theta(\delta)) := (\sin(\theta(\delta)), 0, \cos(\theta(\delta)))^T \text{ with } \theta(\delta) := \arcsin\left(\frac{\delta}{\sqrt{\delta^2 + (\delta-1)^2}}\right).$$

We are then able to compute another approximation to $\partial_e \mathcal{E}(0)$ using a difference quotient

$$DQ_h := \frac{(\mathcal{E}_h(\theta(\delta_h)) - \mathcal{E}_h(-\theta(\delta_h)))}{(\theta(\delta_h) - \theta(-\delta_h))}$$

of the energies. The function θ and the values of δ_h are chosen so that $X_1(\pm\theta(\delta_h))$ lie on a vertex of the grid. The results are tabulated in Table 4.1. Observe that the energy $\mathcal{E}_h(0)$, the difference quotient DQ_h and the derivative $(\partial_e J^*)_h(u_h)$ appear to converge as $h \to 0$. The experimental order of convergence of the derivative quantities are displayed in Table 4.2.

4.3.1.2 Experiment for simple particles lying on vertices

For this experiment, we compute approximations of the energy and the derivative on a sequence of configurations parametrised by the location of one point $X_1(t)$. The configuration is defined

h	$E_{\partial_e J_h^*}$	E_{DQ_h}	$\mathrm{EOC}_{\partial_e J_h^*}$	EOC_{DQ_h}
0.301511	0.165134	0.156597	–	–
0.152499	0.0520672	0.0567013	1.69327	1.49032
0.0764719	0.0110707	0.0179647	2.24306	1.66523
0.0382639	0.00219195	0.00448579	2.33893	2.00384
0.0191355	–	–	–	–

Table 4.2: Derived quantities for experiment in Subsection 4.3.1.1

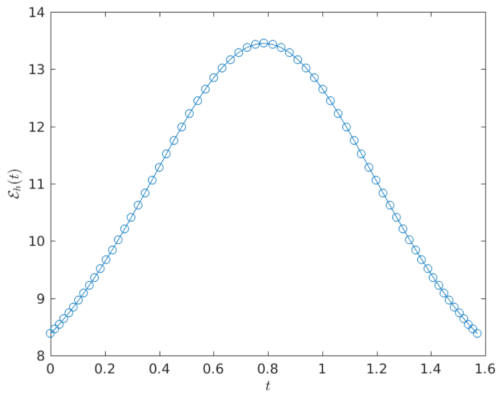


Figure 4.1: Energy $\mathcal{E}_h(t)$ for experiment in Subsection 4.3.1.2

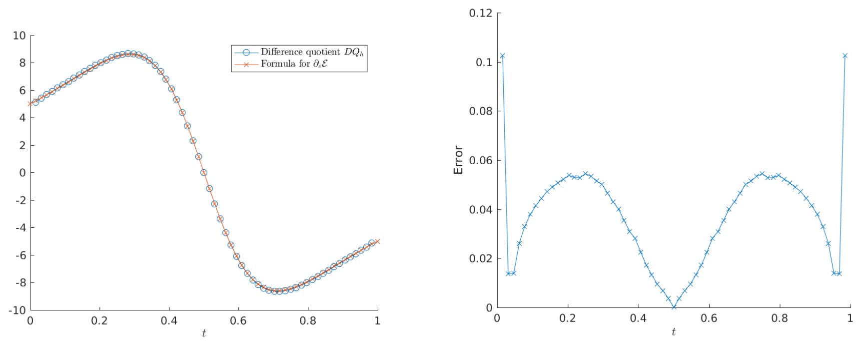
for each t by

$$X_1(t) = (\sin(\theta(t)), 0, \cos(\theta(t)))^T, \;\; Z_1 = 0.1;$$
$$X_2 = (0,0,-1)^T, \;\; Z_2 = 0; \qquad X_3 = (0,1,0)^T, \;\; Z_3 = 0;$$
$$X_4 = (0,-1,0)^T, \;\; Z_4 = 0; \qquad X_5 = (-1,0,0)^T, \;\; Z_5 = 0,$$

where, θ is again defined by $\theta(t) := \arcsin\left(\frac{t}{\sqrt{t^2+(t-1)^2}}\right)$. With this choice of θ we have that the points $X_1, ..., X_5$ lie on vertices of our chosen grid for each evaluation of t. We calculate $\mathcal{E}_h(t)$ and $(\partial_e J^*)_h(u_h(t))$ for $t \in \left\{\frac{m}{2^5} : m \in \mathbb{N}_0, m \leq 2^5\right\}$. In Figure 4.1, we plot $\mathcal{E}_h(t)$. The values $(\partial_e J^*)_h(u_h(t))$ with the difference quotient of $\mathcal{E}_h(t)$ and also the difference between them are given in Figure 4.2. One may calculate that the relative error has a maximum of 2% at the boundary and is below 1% for the interior.

4.3.1.3 Experiment for simple particles not lying on vertices of the grid

We now provide a perturbation of the above experiment. This experiment is to demonstrate that when the constraint points do not lie on the vertices of the grid, the difference quotient becomes a less reliable method. For this experiment we choose $t \in \{\frac{m}{100} : m \in \mathbb{N}_0, m \leq 100\}$. We plot the



(a) $(\partial_e J^*)_h$ and DQ_h (b) $|(\partial_e J^*)_h - DQ_h|$

Figure 4.2: Graphs of quantities from experiment in Subsection 4.3.1.2

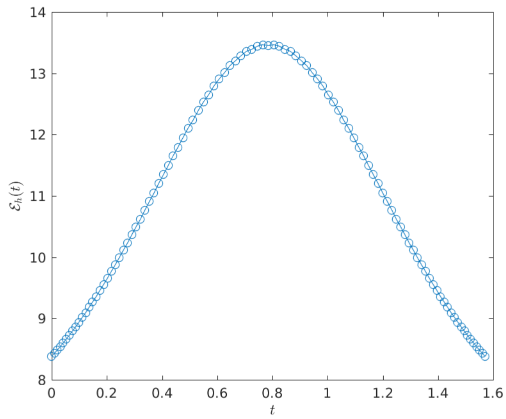
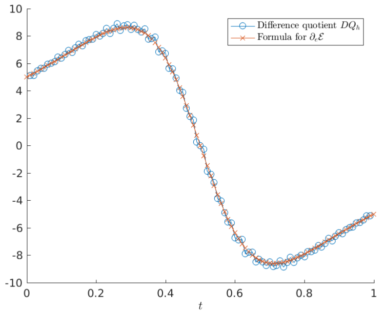


Figure 4.3: Energy $\mathcal{E}_h(t)$ for experiment in Subsection 4.3.1.3

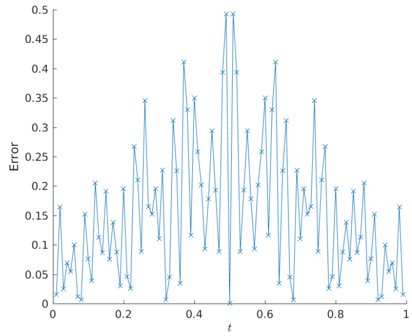
same quantities as in the previous experiment. In Figure 4.3, we plot $\mathcal{E}_h(t)$, we notice it has the same characteristic shape as the previous experiment. For Figure 4.4, we plot $(\partial_e J^*)_h(u_h(t))$ with the difference quotient of $\mathcal{E}_h(t)$ and also the difference between them. We notice that here, the difference quotient does not match the formula as well as in the previous experiment.

4.3.1.4 Experiment for non-trivial particles

This experiment now deals with two non-trivial particles whereby there is little chance of the points lying on vertices unless one is tailoring the grid to the points. We will see that the difference quotients become highly unreliable. We describe the base of the particle $\mathcal{C}_1$ with



(a) $(\partial_e J^*)_h$ and DQ_h



(b) $|(\partial_e J^*)_h - DQ_h|$

Figure 4.4: Graphs of quantities from experiment in Subsection 4.3.1.3

centre $X_{\mathcal{G}_1} = (0,0,1)^T$. We have that $\mathcal{C}_1 = \{X_i\}_{i=1}^8$ with

$$X_1 =(0.5, 0, \sqrt{1-0.5^2})^T, \qquad X_2 =(-0.5, 0, \sqrt{1-0.5^2})^T,$$
$$X_3 =(0.25, 0.25, \sqrt{1-0.25^2-0.25^2})^T, \qquad X_4 =(-0.25, 0.25, \sqrt{1-0.25^2-0.25^2})^T,$$
$$X_5 =(0.25, -0.25, \sqrt{1-0.25^2-0.25^2})^T, \qquad X_6 =(-0.25, -0.25, \sqrt{1-0.25^2-0.25^2})^T,$$
$$X_7 =(0, 0.125, \sqrt{1-0.125^2})^T, \qquad X_8 =(0, -0.125, \sqrt{1-0.125^2})^T,$$

and $(Z_1)_i = 1 - \frac{1}{5}(X_i)_1^2$ for $i = 1, ..., 8$. We let

$$\mathcal{C}_2 := \{x = (x_1, x_2, x_3)^T \in \Gamma : (x_1, x_3, -x_2)^T \in \mathcal{C}_1\},$$

with $(Z_2)_i = 1 - \frac{1}{5}(X_i)_1^2$ for $i = 1, ..., 8$.

We consider the rotation of $\mathcal{C}_1$ about the north pole, we write $\mathcal{C}_1(t) := \mathcal{C}(0, \frac{\pi}{2}t)$. We calculate the quantities $\mathcal{E}_h(t)$ and $(\partial_e J^*)_h(u_h(t))$ for $t \in \{\frac{m}{2^4} : m \in N_0, m \leq 2^5\}$. We plot $\mathcal{E}(t)$ in Figure 4.5. In Figure 4.6 we plot $(\partial_e J^*)_h(u_h(t))$ and the central difference quotient for $\mathcal{E}_h(t)$.

4.3.1.5 Experiment to observe the numerical error of a trivial system

We notice that the difference quotient in the previous experiment is extremely noisy, in this experiment, we consider a perturbation of the above experiment, where we remove $\mathcal{C}_2$ so that, in light of Corollary 4.2.15, we are approximating zero. The quantities from this experiment are plotted in Figure 4.7 where it is seen that there are moderately large perturbations from the average of the energy and the derivative is quite small, as expected.

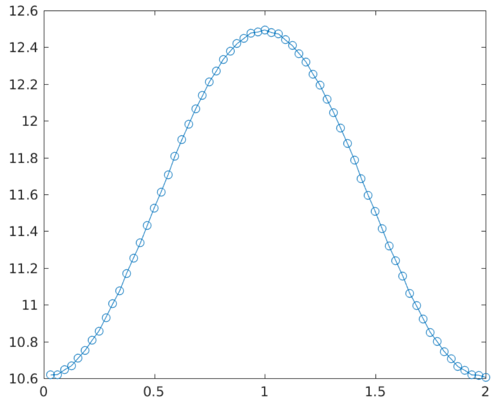
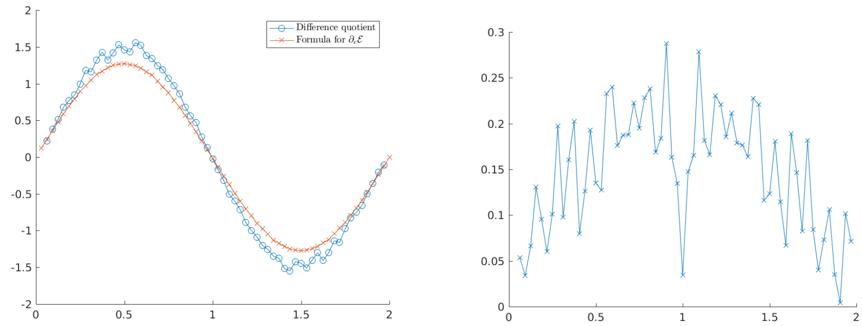


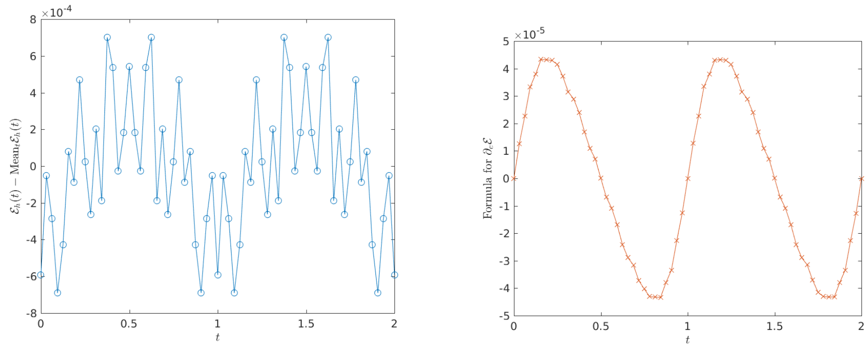
Figure 4.5: Energy $\mathcal{E}_h(t)$ for experiment in Subsection 4.3.1.4



(a) $(\partial_e J^*)_h$ and DQ_h

(b) $|(\partial_e J^*)_h - DQ_h|$

Figure 4.6: Graphs of quantities from experiment in Subsection 4.3.1.4



(a) Deviation from average of $\mathcal{E}_h(t)$

(b) $(\partial_e J^*)_h$

Figure 4.7: Graphs for experiment in Subsection 4.3.1.5

4.3.1.6 Application of formula

We now give the results of a numerical experiment which shows that for a perturbation of our non-trivial particles, they demonstrate a preferential orientation. The idea of our experiment is to consider a particle based at a pole and a particle based at the equator. We then calculate the derivative of the energy as the particle at the pole is moved towards the particle at the equator. This experiment is then redone after rotating the particle at the pole by $\frac{\pi}{2}$. We define the particle $\mathcal{C}_1 = \{X_i\}_{i=1}^8$ by

$$
\begin{aligned}
X_1 &= (0.3, 0, \sqrt{1-0.3^2})^T, & X_2 &= (-0.3, 0, \sqrt{1-0.3^2})^T, \\
X_3 &= (0.15, 0.15, \sqrt{1-0.15^2-0.15^2})^T, & X_4 &= (-0.15, 0.15, \sqrt{1-0.15^2-0.15^2})^T, \\
X_5 &= (0.15, -0.15, \sqrt{1-0.15^2-0.15^2})^T, & X_6 &= (-0.15, -0.15, \sqrt{1-0.15^2-0.15^2})^T, \\
X_7 &= (0, 0.075, \sqrt{1-0.075^2})^T, & X_8 &= (0, -0.075, \sqrt{1-0.075^2})^T,
\end{aligned}
$$

and $(Z_1)_i = 1 - 0.9(X_i)_1^2$ for $i = 1, ..., 8$. We give this centre $X_{\mathcal{G}_1} := (0, 0, 1)^T$. We define $\mathcal{C}_2$ by

$$\mathcal{C}_2 := \{x = (x_1, x_2, x_3)^T \in \Gamma : (x_1, x_3, -x_2)^T \in \mathcal{C}_1\},$$

with $(Z_2)_i = 1 - 10(X_i)_1^2$ for $i = 1, ..., 8$ and centre .

We calculate the derivative at $0 \in \prod_{i=1}^2 (\mathbb{R} \times T_{X_{\mathcal{G}_i}}\Gamma)$ in direction $e = (0, \tau, 0, 0)$, where $\tau = (0, 1, 0)^T \in T_{X_{\mathcal{G}_1}}$ represents the translation of $\mathcal{C}_1$ in the direction τ.

We then calculate the derivative at $p := (\frac{\pi}{2}, 0, 0, 0)$ in the same direction e.

We find that

$$(\partial_e J^*)_h(0) \approx -10.6729 \quad \text{and} \quad (\partial_e J^*)_h(p) \approx 18.5636.$$

This shows that the orientation affects whether the particles are attracted to each other, with one orientation being repulsive and the other attractive. In Figure 4.8 we give the numerical approximations for membranes $u(0)$ and $u(p)$.

4.4 Conclusion

In this chapter we have shown the differentiability of $\mathcal{E}(p)$, the membrane mediated interaction energy for a near spherical membrane with particles attached at points which depend smoothly on p. Further to showing the differentiability, we have given an explicit formula to calculate the derivative and give numerical examples which demonstrate that this formula would appear to be more robust than a difference quotient approach.

It would be of interest to extend this analysis for particles which are able to move more generally, tilting and moving out from the surface. Furthermore it is desirable to consider the problem for inequality constraints on the 'interior' of a particle. Finally, one could analyse higher

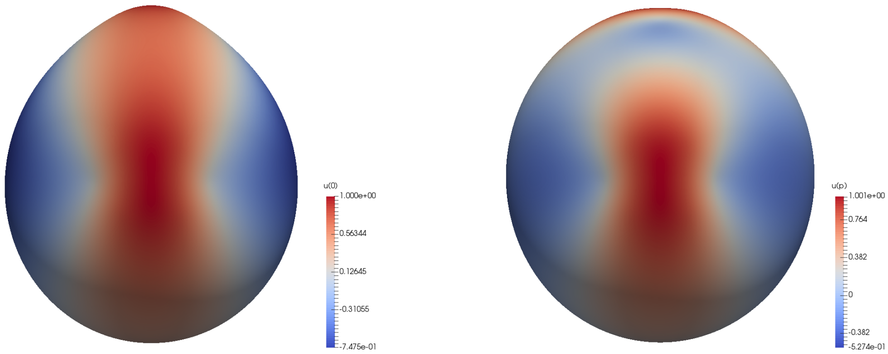


Figure 4.8: The rescaled graphs of the membranes from experiment in Subsection 4.3.1.6, left $0.1\,u(0)$, right $0.1\,u(p)$, both with $(0,1,0)^T$ coming out of the page and $(0,0,1)^T$ pointing up. The colours represent the magnitude of the deformation.

order derivatives of the energy so that one could determine stability of a given configuration.

Chapter 5

Conclusion

In this thesis, we have worked to extend the work of [33] to include a more general problem with constraints, the work of [50] to the energy constructed in [32].

There are many interesting questions related to this thesis worth studying for which we did not have time. We now briefly discuss a few of them.

5.1 Dynamics of a membrane particle system

In Chapter 4 we constructed the derivative of the energy of a membrane particle system. We also gave a construction so that the derivative is easy to calculate. It would be natural to consider a gradient flow or gradient descent using the derivative. That is to say, consider the solution of

$$\partial_t p_t = -\nabla_p \overline{\mathcal{E}}(p_t), \quad p_0 = P_0,$$

or

$$p_{n+1} = p_n - \alpha_n \nabla_p \overline{\mathcal{E}}(p_n), \quad p_0 = P_0,$$

where $\overline{\mathcal{E}}$ is sum of a protein-protein interaction energy and the energy of the membrane $\mathcal{E}$ and $P_0 \in \Lambda^\circ$ is a given initial configuration. It is interesting to consider if such dynamics would mirror that of the natural systems and if it demonstrates the formation of structures similar to those shown in [58, Figure 5.3.1].

It is also worthwhile considering the stochastic variants of the above dynamics. With appropriate noise, the dynamics may be used to estimate statistical quantities of the system, or as a method to help find global minimisers with so-called simulated annealing. Studying the dynamics of the membrane-particle system ties in with the following topic.

5.2 Comparison to real data

We have proposed a method which is meant to model a biomembrane with point constraints, and dynamics to study. An obvious question to ask is, with appropriate tuning of parameters, do these dynamics mirror those of the physical system it is meant to model? A recent article considers this for a different biological problem. In [82], the authors compare a surface phase field model for lipid domain formation to biological lipid domain formation.

5.3 Willmore flow with point constraints

In light of the very recent article [61] which gives a convergent finite element algorithm for the Willmore flow of closed surfaces, a natural question is: how can one extend the developed theory to deal with the point constraints we wish to consider? If a flow were to become near-stationary, it would be interesting to compare how different the almost stationary solution is to the solution produced by small deformation models.

5.4 Differentiability with inequality constraints

As mentioned in the conclusion of Chapter 4, it would be interesting to consider if one could show differentiability of the membrane-particle energy if one were to add in inequality constraints to the model. The inequality constraint would relate to the fact that a protein should stay completely on one side of the membrane. It is clear that there are many different mathematical challenges related to this problem.

5.5 Small deformations of a near-tube membrane

A near-tube membrane model appears in [58] as a membrane energy which is considered for the differentiability with respect to curve constraints. It would be interesting to consider a more rigorous derivation than that which is given, along the lines of [32].

Membrane tubes, also known as membrane tubules, are common structures in many cellular organelles and have many diverse roles within a cell. Tubular structures have been identified as early as the 1950s under examination by electron microscopy [70], where they are seen to be part of the endoplasmic reticulum, a network of interconnected tubules and cisternae. Membrane tubes can be generated by in-vitro tether-pulling experiments [21, 64, 77, 24]. It has been demonstrated that cells exchange enclosed material via the formation of similar narrow fluid membrane tubes known as tunnelling nanotubes [65, 46]. These membrane tubes typically have a diameter of 50 nm to 200 nm and may extend over tens of microns. For further details on the physics of membrane tubes, we refer the reader to [72]. The article [78] discusses how proteins on the membrane are important for tubulation to occur on a membrane and [4] explores the formation and stability of tubes.

Unlike in [32], it is not necessarily reasonable to assume a constraint on the "enclosed volume" as the tube is not a closed surface. In addition, one need not expect that the first variation of a Lagrangian associated to the Canham-Helfrich energy (1.1) would vanish without an extra constraint. As such, we suggest that one should constrain the boundary, the surface area and the integral of the geodesic curvature on the boundary. By using the small deformations methodology around

$$\Gamma := \left\{(x_1, x_2, x_3)^T \in \mathbb{R}^3 : x_1^2 + x_2^2 = R^2,\, x_3 \in (-L/2, L/2)\right\},$$

the tube of length $L > 0$ and radius $R > 0$, one arrives at the following small deformations Lagrangian

$$\mathcal{L}_{\text{Tube}}(u, \mu) := \kappa \int_\Gamma (\Delta_\Gamma u)^2 - \frac{2}{R^2}(\tau \cdot \nabla_\Gamma u)^2 + \frac{1}{R^4}u^2 + \mu_1 \int_\Gamma u + \mu_2 \int_\Gamma \Delta_\Gamma u - \langle f, u\rangle,$$

where $\tau(x_1, x_2, x_3) := \frac{(x_2, -x_1, 0)^T}{R}$ is a unit tangential vector field in the azimuthal direction and $f \in \left(H^2(\Gamma) \cap H_0^1(\Gamma)\right)^*$ is an external forcing term. The development of this small deformations Lagrangian requires the calculation of the second variation of the geometric terms for surfaces with boundary. This differs from the calculations of [32] as the fact that the first variation of the first variation coincides with the second variation for closed surfaces is exploited.

It is possible to verify that the bilinear form of this Lagrangian is coercive over $H^2(\Gamma) \cap H_0^1(\Gamma)$ with the non-standard norm

$$\|u\|_{2,2} := \left(\int_\Gamma (\Delta_\Gamma u)^2 + |\nabla_\Gamma u|^2 + u^2\right)^{1/2},$$

where the coercivity constant depends on the aspect ratio of the tube. It is important to verify that this norm is equivalent to the standard norm on $H^2(\Gamma) \cap H_0^1(\Gamma)$.

It is interesting to consider the derivative of this membrane energy with respect to point constraints, as done in Chapter 4. One would expect that the theory and calculations should follow almost identically.

It is also important to note that one may not simply apply the abstract theory presented in Chapter 3 for existence and uniqueness or the approximation error of a split problem. This is because there is a constraint on $\Delta_\Gamma u$. It is possible to show well-posedness to a split problem, where one sets $w := -\Delta_\Gamma u + \mu_2$, where μ_2 is the Lagrange multiplier associated to constraining the integral of $\Delta_\Gamma u$. It is also possible to show that the appropriate discrete problem is well-posed.

5.5.1 Numerical experiments for a membrane tube problem

We now provide some numerical experiments related to the critical points of $\mathcal{L}_{\text{Tube}}$ with non-homogeneous boundary terms and a non-trivial forcing f. For simplicity, we set $R = 1$ and $L = 2$ and $\kappa = 1$.

5.5.1.1 Experiment for the approximation of an axi-symmetric membrane with inhomogeneous boundary

We begin by considering an axi-symmetric example. The problem we consider is to find (u, μ_1, μ_2) such that:

$$
\begin{aligned}
\Delta_\Gamma^2 u + 2\text{div}_\Gamma\left(\tau \otimes \tau \nabla_\Gamma u\right) + u =& -\mu_1 \quad \text{ in } \Gamma,\\
u|_{\partial\Gamma} =& 1,\\
\Delta_\Gamma u|_{\partial\Gamma} =& \mu_2,\\
\int_\Gamma u =& 0,\\
\int_\Gamma \Delta_\Gamma u =& 0,
\end{aligned}
$$

with splitting

$$
\begin{aligned}
-\Delta_\Gamma w + 2\text{div}_\Gamma\left(\tau \otimes \tau \nabla_\Gamma u\right) + u = -\mu_1 \text{ in } \Gamma,\\
-\Delta_\Gamma u - w + \mu_2 = 0 \text{ in } \Gamma,\\
\int_\Gamma u = \int_\Gamma (w - \mu_2) = 0,\\
u|_{\partial\Gamma} = 1,\\
w|_{\partial\Gamma} = 0.
\end{aligned}
$$

By axi-symmetry of the data, it is clear that the solution u will be a linear combination of $\{\varphi_j\}_{j=1}^4$, where

$$
\begin{aligned}
\varphi_1(z) =& \sin\left(\frac{z}{\sqrt{2}}\right)\sinh\left(\frac{z}{\sqrt{2}}\right), \quad \varphi_2(z) = \cos\left(\frac{z}{\sqrt{2}}\right)\sinh\left(\frac{z}{\sqrt{2}}\right),\\
\varphi_3(z) =& \sin\left(\frac{z}{\sqrt{2}}\right)\cosh\left(\frac{z}{\sqrt{2}}\right), \quad \varphi_4(z) = \cos\left(\frac{z}{\sqrt{2}}\right)\cosh\left(\frac{z}{\sqrt{2}}\right).
\end{aligned}
$$

along with a correction to constrain the mean value.

We see that u, w, μ_1 and μ_2, the continuous solutions, are given by

$$\begin{aligned} u(x_1,x_2,x_3) &= \sum_{i=1}^{4} \beta_i \left(\varphi_i(x_3) - \frac{1}{2}\int_{-1}^{1} \varphi_i \right), \\ \mu_1 &= -\frac{1}{2}\sum_{i=1}^{4} \beta_i \int_{-1}^{1} \varphi_i, \\ \mu_2 &= \sum_{i=1}^{4} \beta_i \varphi_i''(1), \\ w(x_1,x_2,x_3) &= \mu_2 - \sum_{i=1}^{4} \beta_i \varphi_i''(x_3), \end{aligned} \tag{5.1}$$

where $\beta \approx (7.4792, 0, 0, 44.55442)^T$ is determined by solving an appropriate (4-dimensional) linear system. The linear system we consider, is given as

$$\begin{aligned} \sum_{i=1}^{4} \int_{-1}^{1} \beta_i \varphi_i'' &= 0, \\ \sum_{i=1}^{4} \beta_i (\varphi_i''(-1) - \varphi_i''(1)) &= 0, \\ \sum_{i=1}^{4} \beta_i \left(\varphi_i(1) - \frac{1}{2}\int_{-1}^{1} \varphi_i \right) &= 1, \\ \sum_{i=1}^{4} \beta_i \left(\varphi_i(-1) - \frac{1}{2}\int_{-1}^{1} \varphi_i \right) &= 1. \end{aligned}$$

We also calculate a numerical approximation of the solution of this problem by surface finite elements. We compare this to the interpolation of the solution in (5.1). We choose to estimate the w error in the H^1 norm because, by smoothness of our data, we might expect it to converge with order h.

The errors calculated with quadrature are displayed in Tables 5.1, 5.2 and 5.3.

h	E_{L^2}	E_{H^1}	EOC_{L^2}	EOC_{H^1}
0.287404	0.122095	1.66157	-1	-1
0.144767	0.0315451	0.854108	1.97355	0.970399
0.0727011	0.00795161	0.429965	2.00073	0.996489
0.0364239	0.00199388	0.215347	2.00149	1.00047
0.0182314	0.00049819	0.107719	2.00244	1.00093

Table 5.1: Errors and experimental order of convergence for $u - u_h^l$ in the axi-symmetric experiment, Section 5.5.1.1.

h	E_{L^2}	E_{H^1}	EOC_{L^2}	EOC_{H^1}
0.287404	0.199395	11.7842	-1	-1
0.144767	0.0499988	5.89657	2.01716	1.00966
0.0727011	0.0125108	2.94883	2.01142	1.00608
0.0364239	0.00314142	1.47449	2.0055	1.00285
0.0182314	0.000782015	0.737251	2.00325	1.00152

Table 5.2: Errors and experimental order of convergence for $w - w_h^l$ in the axi-symmetric experiment, Section 5.5.1.1.

h	E_{μ_1}	E_{μ_2}	EOC_{μ_1}	EOC_{μ_2}
0.287404	0.0729122	0.248456	-1	-1
0.144767	0.0210297	0.0630323	1.81305	2.00014
0.0727011	0.00545994	0.0158357	1.95783	2.00557
0.0364239	0.00140207	0.00398226	1.96705	1.99733
0.0182314	0.000373319	0.00101422	1.91202	1.97626

Table 5.3: Errors and experimental orders of convergence for $\mu_1 - \mu_1^h$ and $\mu_2 - \mu_2^h$ in the axi-symmetric experiment, Section 5.5.1.1.

5.5.1.2 Experiment for the approximation of a membrane with point contraints

For this example, we consider the problem with point constraints to find $(u, \mu_1, \mu_2, \lambda)$ such that:

$$
\begin{aligned}
-\Delta_\Gamma w + 2\mathrm{div}_\Gamma\left(\tau \otimes \tau \nabla_\Gamma u\right) + u =& f - \mu_1 \quad \text{ in } \Gamma \setminus \{X_1, X_2, X_3, X_4\},\\
-\Delta_\Gamma u - w + \mu_2 =& 0 \quad \text{ in } \Gamma,\\
u|_{\partial\Gamma} =& U|_{\partial\Gamma},\\
w|_{\partial\Gamma} =& W|_{\partial\Gamma},\\
\int_\Gamma u =& \alpha,\\
\int_\Gamma w - \mu_2 =& \beta,\\
u(X_i) =& Z_i \text{ for } i = 1, 2, 3, 4,
\end{aligned}
$$

where

$$
\begin{aligned}
U =&(2-2x_2+x_3^2)\log(2-2x_2+x_3^2)-2(2-2x_2+x_3^2),\\
W =&\frac{1}{\frac{1}{x_3^2}+\frac{1}{2-2x_2}}+2(1+x_2)\log(2-2x_2+x_3^2),\\
f =&\frac{2x_3^8}{(2-2x_2+x_3^2)^3}-\frac{2x_3^4(x_3^2-8)}{(2-2x_2+x_3^2)^2}-\frac{x_3^2(x_3^2+16)}{2-2x_2+x_3^2}+(2+x_3^2)(-1+\log(2-2x_2+x_3^2)),\\
X_1 =&(0,1,0)^T, \quad X_2=(0,-1,0)^T, \quad X_3=(1,0,0)^T, \quad X_4=(-1,0,0)^T,\\
g =&0, \quad \alpha=\int_\Gamma U, \quad \beta=\int_\Gamma W, \quad Z_i=U(X_i) \text{ for } i=1,2,3,4.
\end{aligned}
$$

We see that $(u,w,\mu_1,\mu_2,\lambda)$ which solves the above system weakly is given by

$$u=U, \quad w=W, \quad \mu_1=0, \quad \mu_2=0, \quad \lambda=(-4\pi,0,0,0)^T.$$

The errors calculated with quadrature between the interpolated exact solution and the surface finite element solution are displayed in Tables 5.4, 5.5 and 5.6. We see that in Table 5.5, the error in $W^{1,4/3}$ is rather large. Using the finest grid, we calculate $\|w\|_{1,4/3}$ to be approximately 155.205, as such, the relative error on the finest grid is roughly 2.23%.

h	E_{L^2}	E_{H^1}	EOC_{L^2}	EOC_{H^1}
0.287404	0.0230088	0.686905	–	–
0.144767	0.00547454	0.349979	2.09369	0.983318
0.0727011	0.00131694	0.175994	2.0686	0.998042
0.0364239	0.000322465	0.0881458	2.0359	1.00047
0.0182314	0.0000806517	0.0440943	2.00245	1.00084

Table 5.4: Errors and experimental order of convergence for $u-u_h^l$ in the point constraint experiment, Subsection 5.5.1.2.

h	E_{L^2}	$E_{W^{1,4/3}}$	EOC_{L^2}	$EOC_{W^{1,4/3}}$
0.287404	1.66166	13.6087	–	–
0.144767	0.660204	9.51081	1.34598	0.522455
0.0727011	0.270766	6.78274	1.29404	0.490799
0.0364239	0.120391	4.84791	1.17273	0.485918
0.0182314	0.0571026	3.45644	1.07776	0.488829

Table 5.5: Errors and experimental order of convergence for $w-w_h^l$ in the point constraint experiment, Subsection 5.5.1.2.

Theory related to this membrane tube energy will be addressed in a work in preparation.

h	E_λ	E_{μ_1}	E_{μ_2}	EOC_λ	EOC_{μ_1}	EOC_{μ_2}
0.287404	0.242688	0.188594	0.193561	–	–	–
0.144767	0.133595	0.0661003	0.030098	0.870513	1.52884	2.71397
0.0727011	0.0501027	0.0210878	0.00615303	1.42389	1.65872	2.30485
0.0364239	0.0165791	0.00639441	0.00204711	1.60018	1.72655	1.59234
0.0182314	0.0051533	0.00187797	0.000722672	1.68839	1.77036	1.50449

Table 5.6: Errors and experimental orders of convergence for $\lambda - \lambda_h$, $\mu_1 - \mu_1^h$ and $\mu_2 - \mu_2^h$ in the point constraint experiment, Subsection 5.5.1.2.

5.6 Weaker assumptions in Chapter 3

In Chapter 3 we required some strong assumptions. Two assumptions in particular are obvious places to relax, that S is a Hilbert space and that the map $T\colon X \to S$ is surjective. These two restrictions are the main barrier to the choices of problems one can easily study. We give the following example: consider c, b and m as in Section 3.3.1 with a line constraint on some curve γ, that is to say T is the trace operator $T\colon u \mapsto u|_\gamma$. Supposing that we had irregular data, $f \in \left(W^{1,q}(\Gamma)\right)^*$ for $q > 2$ so that we should still set X to be functions in $W^{1,q}(\Gamma)$ with vanishing integral. With this choice of X, if we were to require T to be surjective, this means we would need S to be $W^{1-\frac{1}{q},q}(\gamma)$. Penalising the square of the $W^{1-\frac{1}{q},q}(\gamma)$ norm could be troublesome from the perspective of numerical implementation, both from the non-integer derivative and the non-linearity. As such, it would be convenient if one did not require that T was surjective.

Studying a non-linear second order splitting would also be of interest, particularly with respect to the penalty formulation. This penalty formulation may have non-linearities introduced by penalising non-Hilbert norms as mentioned above, or potentially by considering the penalisation of convex conditions e.g. for a penalty formulation of obstacle-type problems.

Appendix A

The pullback to a reference domain

We give some general results on the calculation of the composition of pullbacks and derivatives, where we consider that the image and domain of the diffeomorphism need not be the same. As we are working with different surfaces, we will need to make clear to which surface geometric quantities belong to, this is done with a superscript of the surface, e.g. H^{Γ_1} is the mean curvature of Γ_1 and H^{Γ_0} the mean curvature of Γ_0. Consider the case of Γ_0 and Γ_1 being C^k, compact surfaces, with $\mathrm{X}\colon \Gamma_0 \to \Gamma_1$ a C^k-diffeomorphism, where we require $k \geq 2$.

Given some function $u\colon \Gamma_1 \to \mathbb{R}$ we wish to obtain expressions for $(\nabla_{\Gamma_1} u) \circ \mathrm{X}$ and $(\underline{D}^2_{\Gamma_1} u) \circ \mathrm{X}$. The first part of this is developed in [18], where also the trace of the second quantity, the Laplace-Beltrami, is calculated. Although for the model we consider in this work, the surface Hessian is not required, we compute it for completion as it may arise in other elastic type models, where the Hessian regularly arises. We choose to do this in an method which avoids integration by parts so that surfaces with boundary may be considered.

Lemma A.0.1. *Let $u \in H^1(\Gamma_1)$, then $u \circ \mathrm{X} \in H^1(\Gamma_0)$ and*

$$(\nabla_{\Gamma_1} u) \circ \mathrm{X} = \left(\nabla_{\Gamma_0} \mathrm{X} + \nu^{\Gamma_1} \circ \mathrm{X} \otimes \nu^{\Gamma_0}\right)^{-T} \nabla_{\Gamma_0}(u \circ \mathrm{X}) = \nabla_{\Gamma_0} \mathrm{X} G_{\Gamma_0}^{-1} \nabla_{\Gamma_0}(u \circ \mathrm{X}),$$

where $G_{\Gamma_0} := \nabla_{\Gamma_0} \mathrm{X}^T \nabla_{\Gamma_0} \mathrm{X} + \nu^{\Gamma_0} \otimes \nu^{\Gamma_0}$.

The proof is shown in Lemma 3.2 of [18]. We write $B := \nabla_{\Gamma_0} \mathrm{X} + \nu^{\Gamma_1} \circ \mathrm{X} \otimes \nu^{\Gamma_0}$, which satisfies

$$B^T B = G_{\Gamma_0}.$$

This gives a simpler form of the above lemma,

$$(\nabla_{\Gamma_1} u) \circ \mathrm{X} = B^{-T} \nabla_{\Gamma_0}(u \circ \mathrm{X}).$$

Lemma A.0.2. *Let $u \in H^2(\Gamma_1)$, then $u \circ \mathrm{X} \in H^2(\Gamma_0)$ and for $i, j = 1, ..., n+1$*

$$\begin{aligned}(\underline{D}_i^{\Gamma_1} \underline{D}_j^{\Gamma_1} u) \circ \mathrm{X} =& \frac{1}{b} \mathrm{div}_{\Gamma_0} \left(b B^{-1} (B^{-T} \nabla_{\Gamma_0} \hat{u})_j \right)_i \\ &+ (H^{\Gamma_1} \circ \mathrm{X} - H^{\Gamma_0})(\nu_i^{\Gamma_1} \circ \mathrm{X})(B^{-T} \nabla_{\Gamma_0} \hat{u})_j,\end{aligned}$$

where $b = \det(B)$, $b_{ij} = B_{ij}$ and $b^{ij} = \left(B^{-1}\right)_{ij}$.

Proof. We write $\hat{u} := u \circ \mathrm{X}$ and where indices are repeated in a product, summation is assumed. We now make use of the preceding lemma to obtain,

$$\underline{D}_i^{\Gamma_1} \underline{D}_j^{\Gamma_1} u \circ \mathrm{X} = b^{li} \underline{D}_l^{\Gamma_0} \left(b^{kj} \underline{D}_k^{\Gamma_0} \hat{u} \right).$$

We now put this into something similar to a divergence form,

$$\underline{D}_i^{\Gamma_1} \underline{D}_j^{\Gamma_1} u \circ \mathrm{X} = = \quad \frac{1}{b} \underline{D}_l^{\Gamma_0} \left(b b^{li} b^{kj} \underline{D}_k^{\Gamma_0} \hat{u} \right) - \frac{1}{b} \underline{D}_l^{\Gamma_0} (b)\, b^{li} b^{kj} \underline{D}_k^{\Gamma_0} \hat{u} - \underline{D}_l^{\Gamma_0} \left(b^{li} \right) b^{kj} \underline{D}_k^{\Gamma_0} \hat{u}.$$

In [18], it is calculated

$$\underline{D}_l^{\Gamma_0} b^{li} = -b^{lm} \underline{D}_l^{\Gamma_0} b_{mf} b^{fi}, \quad \frac{1}{b} \underline{D}_l^{\Gamma_0} b = b^{fm} \underline{D}_l^{\Gamma_0} b_{mf},$$

inserting these into the above gives,

$$\underline{D}_i^{\Gamma_1} \underline{D}_j^{\Gamma_1} u \circ \mathrm{X} = \frac{1}{b} \underline{D}_l^{\Gamma_0} \left(b b^{li} b^{kj} \underline{D}_k^{\Gamma_0} \hat{u} \right) - b^{fm} \underline{D}_l^{\Gamma_0} b_{mf} b^{li} b^{kj} \underline{D}_k^{\Gamma_0} \hat{u} + b^{lm} \underline{D}_l^{\Gamma_0} b_{mf} b^{fi} b^{kj} \underline{D}_k^{\Gamma_0} \hat{u}.$$

Since we are summing over f, k, l and m in the above, it is possible to swap the indices, in particular we swap f and l in the second term. We now consider the terms

$$\begin{aligned}b^{lm} \underline{D}_l^{\Gamma_0} b_{mf} b^{fi} b^{kj} \underline{D}_k^{\Gamma_0} \hat{u} - & b^{lm} \underline{D}_f^{\Gamma_0} b_{ml} b^{fi} b^{kj} \underline{D}_k^{\Gamma_0} \hat{u} \\ &= b^{lm} b^{fi} \left(b^{kj} \underline{D}_k^{\Gamma_0} \hat{u} \right) \left(\underline{D}_l^{\Gamma_0} b_{mf} - \underline{D}_f^{\Gamma_0} b_{ml} \right).\end{aligned} \tag{A.1}$$

In order to simplify this, we will use the definition of B and swap the order of derivatives. As in [18], one calculates

$$\begin{aligned}\underline{D}_l^{\Gamma_0} b_{mf} - \underline{D}_f^{\Gamma_0} b_{ml} = & \left(\underline{D}_l^{\Gamma_0} (\nu_m^{\Gamma_1} \circ \mathrm{X}) - \left(\mathcal{H}^{\Gamma_0} \nabla_{\Gamma_0} \mathrm{X}_m \right)_l \right) \nu_f^{\Gamma_0} \\ &+ \left(\left(\mathcal{H}^{\Gamma_0} \nabla_{\Gamma_0} \mathrm{X}_m \right)_f - \underline{D}_f^{\Gamma_0} (\nu_m^{\Gamma_1} \circ \mathrm{X}) \right) \nu_l^{\Gamma_0}.\end{aligned}$$

We now use this to simplify (A.1). We will make use of the relation $b^{ki} \nu_k^{\Gamma_0} = \nu_i^{\Gamma_1} \circ \mathrm{X}$. We calculate each part,

$$\begin{aligned}
b^{lm}b^{fi}\underline{D}_l^{\Gamma_0}(\nu_m^{\Gamma_1}\circ \mathrm{X})\nu_f^{\Gamma_0} =&(\nu_i^{\Gamma_1}\circ \mathrm{X})(B^{-T}(\nabla_{\Gamma_1}\nu_m)\circ \mathrm{X})_m\\
=&(H^{\Gamma_1}\nu_i^{\Gamma_1})\circ \mathrm{X},
\end{aligned}$$

$$\begin{aligned}
b^{lm}b^{fi}(\mathcal{H}^{\Gamma_0}\nabla_{\Gamma_0}\mathrm{X}_m)_l\nu_f^{\Gamma_0} =&b^{lm}b^{fi}\mathcal{H}_{lk}^{\Gamma_0}\underline{D}_k^{\Gamma_0}\mathrm{X}_m\nu_f^{\Gamma_0}\\
=&b^{lm}b^{fi}\mathcal{H}_{lk}^{\Gamma_0}b_{mk}\nu_f^{\Gamma_0}\\
=&H^{\Gamma_0}(\nu_i^{\Gamma_1}\circ \mathrm{X}),
\end{aligned}$$

$$\begin{aligned}
b^{lm}b^{fi}(\mathcal{H}^{\Gamma_0}\nabla_{\Gamma_0}\mathrm{X}_m)_f\nu_l^{\Gamma_0} =&b^{lm}b^{fi}\mathcal{H}_{fk}^{\Gamma_0}\underline{D}_k^{\Gamma_0}\mathrm{X}_m\nu_l^{\Gamma_0}\\
=&b^{lm}b^{fi}\mathcal{H}_{fk}^{\Gamma_0}b_{mk}\nu_l^{\Gamma_0}\\
=&b^{fi}\mathcal{H}_{fl}^{\Gamma_0}\nu_l^{\Gamma_0}=0,
\end{aligned}$$

$$\begin{aligned}
b^{lm}b^{fi}(\nu_m^{\Gamma_1}\circ \mathrm{X})\nu_l^{\Gamma_1} =&(\nu_m^{\Gamma_1}\circ \mathrm{X})(B^{-T}\nabla_{\Gamma_0}(\nu^{\Gamma_1}\circ \mathrm{X}))_i\\
=&(\nu_m^{\Gamma_1}\circ \mathrm{X})\mathcal{H}_{mi}^{\Gamma_1}\circ \mathrm{X}=0.
\end{aligned}$$

This then gives

$$b^{lm}b^{fi}\left(b^{kj}\underline{D}_k^{\Gamma_0}\hat{u}\right)\left(\underline{D}_l^{\Gamma_0}b_{mf}-\underline{D}_f^{\Gamma_0}b_{ml}\right)=((H^{\Gamma_1}\circ \mathrm{X})-H^{\Gamma_0})(\nu_i^{\Gamma_1}\circ \mathrm{X})(B^{-T}\nabla_{\Gamma_0}\hat{u})_j,$$

which completes the result. □

Remark A.0.3. *By taking the trace of $\underline{D}_{\Gamma_1}^2 u\circ \mathrm{X}$, one obtains*

$$(\Delta_{\Gamma_1}u)\circ \mathrm{X}=\frac{1}{b}div_{\Gamma_0}(bG_{\Gamma_0}^{-1}\nabla_{\Gamma_0}(u\circ \mathrm{X})).$$

Appendix B

Implicit function theorem

We give the version of the implicit function theorem we use in Theorem 4.2.9. The result is taken from [19, Theorem 7.13-1].

Theorem B.0.1. *Let $\mathcal{X}$ be a normed vector space, $\mathcal{Y}$ and $\mathcal{Z}$ Banach spaces with $\Omega \subset \mathcal{X} \times \mathcal{Y}$ open with $(a,b) \in \Omega$. Let $f \in C(\Omega; \mathcal{Z})$ with $f(a,b) = 0$, $\frac{\partial f}{\partial y}(x,y) \in \mathcal{L}(\mathcal{Y}; \mathcal{Z})$ exists at all points $(x,y) \in \Omega$ and $\frac{\partial f}{\partial y} \in C(\Omega; \mathcal{L}(\mathcal{Y}; \mathcal{Z}))$, $\frac{\partial f}{\partial y}(a,b)$ is a bijection, so that $\left(\frac{\partial f}{\partial y}(a,b)\right)^{-1} \in \mathcal{L}(\mathcal{Z}; \mathcal{Y})$.*

1. *Then there is an open neighbourhood V of a in $\mathcal{X}$, a neighbourhood W of b in $\mathcal{Y}$ and $g \in C(V; W)$ such that $V \times W \subset \Omega$ and $\{(x,y) \in V \times W : f(x,y) = 0\} = \{(x,y) \in V \times W : y = g(x)\}$.*

2. *Assume in addition that f is differentiable at $(a,b) \in \Omega$. Then g is differentiable at a and*

$$g'(a) = -\left(\frac{\partial f}{\partial y}(a,b)\right)^{-1} \frac{\partial f}{\partial x}(a,b) \in \mathcal{L}(\mathcal{X}; \mathcal{Y}).$$

3. *Assume in addition that $f \in C^k(\Omega; \mathcal{Z})$ for some $k \geq 1$. Then there is an open neighbourhood $\tilde{V} \subset V$ of a in $\mathcal{X}$ and neighbourhood $\tilde{W} \subset W$ of b in $\mathcal{Y}$ such that $\frac{\partial f}{\partial y}(x,y) \in \mathcal{L}(\mathcal{Y}; \mathcal{Z})$ is a bijection, so that $\left(\frac{\partial f}{\partial y}(x,y)\right)^{-1} \in \mathcal{L}(\mathcal{Z}; \mathcal{Y})$ at each $(x,y) \in \tilde{V} \times \tilde{W}$, $g \in C^k(\tilde{V}; \mathcal{Y})$, $g'(x) = -\left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1} \frac{\partial f}{\partial x}(x, g(x)) \in \mathcal{L}(\mathcal{X}; \mathcal{Y})$ for each $x \in \tilde{V}$.*

Appendix C

Elliptic regularity

We first show, for arbitrary surfaces, that $\Delta_\Gamma u \in W^{1,p}(\Gamma)$ for $p \leq 2$ gives $u \in W^{3,p}(\Gamma)$.

Proposition C.0.1. *Suppose $u \in H^1(\Gamma)$ with $\Delta_\Gamma u \in W^{1,p}(\Gamma)$ for some $p \in (1, 2]$ and Γ is C^3, then there is a $C > 0$ independent of u such that for each $i, j = 1, 2, 3$,*

$$\|\underline{D}_i \underline{D}_j u\|_{1,p} \leq C \left(\|\underline{D}_j \Delta_\Gamma u\|_{0,p} + \|\Delta_\Gamma u\|_{0,2} + \|\nabla_\Gamma u\|_{0,2} \right).$$

Proof. We make use of the following inf-sup condition, shown in [33]:

$$\exists \gamma > 0 : \gamma \|\xi\|_{1,p} \leq \sup_{\eta \in W^{1,q}(\Gamma)} \frac{\int_\Gamma \nabla_\Gamma \eta \cdot \nabla_\Gamma \xi + \eta \xi}{\|\eta\|_{1,q}} \quad \forall \xi \in W^{1,p}(\Gamma).$$

By the fact that Γ has finite measure, it holds that $\|\underline{D}_i \underline{D}_j u\|_{0,p} \leq C \|\underline{D}_i \underline{D}_j u\|_{0,2}$ which we know is controlled by $\|\Delta_\Gamma u\|_{0,2} + \sqrt{\|\mathcal{H}H - 2\mathcal{H}^2\|_{0,\infty}} \|\nabla_\Gamma u\|_{0,2}$, [28]. It is then sufficient to show that $\int_\Gamma \nabla_\Gamma \underline{D}_i \underline{D}_j u \cdot \nabla_\Gamma \eta$ is bounded appropriately. One may calculate

$$\begin{aligned}
\int_\Gamma \nabla_\Gamma \underline{D}_i \underline{D}_j u \cdot \nabla_\Gamma \eta = \int_\Gamma & \underline{D}_j \Delta_\Gamma u \underline{D}_i \eta \\
& + \Big(\left(\mathcal{H} \nabla_\Gamma \underline{D}_k u\right)_j \nu_k - \left(\mathcal{H} \nabla_\Gamma \underline{D}_k u\right)_k \nu_i - \underline{D}_k \left[\left(\mathcal{H} \nabla_\Gamma u\right)_k \nu_j \right] \Big) \underline{D}_i \eta \\
& - \underline{D}_k \underline{D}_j u \left(\mathcal{H} \nabla_\Gamma \eta\right)_k \nu_i - \left(\mathcal{H} \nabla_\Gamma \underline{D}_j u\right)_k \nu_i \underline{D}_k \eta.
\end{aligned}$$

This follows from repeatedly applying integration by parts and swapping the order of derivatives. Applying Hölder's inequality, the result immediately follows. □

Proposition C.0.2. *Let $u \in H^2(\Gamma)$ be the unique solution of Problem 2.1.3, then it holds that for any $p < 2$, $u \in W^{3,p}(\Gamma)$.*

Proof. By [42, Theorem 2.34] and the arguments presented in [38, Section 5], it is clear that there is $\overline{p} \in \mathbb{R}$ and $\lambda \in \mathbb{R}^K$ such that

$$a(u,v) + \overline{p} \int_\Gamma v + \lambda \cdot v|_{\mathcal{C}} = 0 \quad \forall v \in H^2(\Gamma).$$

Let $\eta := -\Delta_\Gamma u - \frac{2}{R^2} u \in L^2(\Gamma)$, then for any $v \in H^2(\Gamma)$,

$$a(u,v) = \int_\Gamma (-\kappa \Delta_\Gamma v + \sigma v)\eta = -\lambda \cdot v|_{\mathcal{C}} - \overline{p} \int_\Gamma v.$$

Let $\phi \in C^\infty(\Gamma)$ and consider the inverse Laplace type map $G\colon L^2(\Gamma) \to H^2(\Gamma)$ such that $G\colon \phi \mapsto v$ where $-\kappa\Delta_\Gamma v + \sigma v = \phi$. Via a local argument, it may be seen that for any $q > 2$, $\|v\|_{0,\infty} \leq C\|\phi\|_{-1,q}$ [69]. Hence

$$\begin{aligned}
\langle \phi, \eta \rangle = \int_\Gamma \phi\eta &= \int_\Gamma (-\kappa\Delta_\Gamma v + \sigma v)\eta \\
&= -\lambda \cdot v|_{\mathcal{C}} - \overline{p} \int_\Gamma v \\
&\leq \|\lambda\|_{\mathbb{R}^M} \|v\|_{0,\infty} + |\overline{p}| \|v\|_{0,1} \\
&\leq C\|\phi\|_{-1,q}.
\end{aligned}$$

Thus we have shown that η represents a bounded linear operator on $W^{-1,q}(\Gamma)$, thus we have shown that $-\Delta_\Gamma u - \frac{2}{R^2} u \in W^{1,q^*}(\Gamma)$. In particular, by Proposition C.0.1, it holds that $u \in W^{3,q^*}(\Gamma)$. Since $q^* < 2$ is arbitrary, the result is complete. □

Appendix D

Transformations satisfying Assumption 4.2.2

Here, we verify Assumption 4.2.2 by constructing χ.

D.0.0.1 Rotation of a single particle

This example pertains to a simple rotation. The example we consider is rotating a single particle whose centre $X_{\mathcal{G}}$ is taken to be the North pole $N := (0, 0, R)^T$, without loss of generality. The points of the particle are contained in the set $B_r(N) := \{x : x_3 > R - r\}$ around the north pole and all other points are contained in the set $B_{r+\epsilon}(N)^C := \{x : x_3 < R - r - \epsilon\}$.

Since this is a 1-parameter family of transformations, we write, with an abuse of notation $\chi(\alpha, \cdot) = \chi(q, \cdot)$ for the diffeomorphism.

We may then explicitly write

$$\chi(\alpha, x) = \eta(x) \left((0, 0, x_3)^T + \cos(\alpha) \left(\frac{N}{R} \times x \right) \times \frac{N}{R} + \sin(\alpha) \left(\frac{N}{R} \times x \right) \right) + (1 - \eta(x)))x,$$

where $\eta \colon \Gamma \to \mathbb{R}$ is a C^k-smooth cut off function such that $\eta = 1$ on $B_r(N)$ and $\eta = 0$ on $B_{r+\epsilon}(N)^C$ and depends only on x_3. It is clear that this χ is smooth with $\chi(\alpha, \cdot)$ having inverse $\chi(-\alpha, \cdot)$ and that it moves the points of the particle based at the north pole as required, while others remain stationary. Furthermore, for each fixed x_3 it ,essentially, is a 2-dimensional rotation about $(0, 0, x_3)$ so the volume element induced by χ is constantly equal to 1.

It is convenient to calculate, for $e = (1, 0)$, $\partial_e \chi(0, x)$,

$$\partial_e \chi(0, x) = \partial_s \left(\chi(s, x) \right) |_{s=0} = \eta(x) \left(\frac{N}{R} \times x \right).$$

One may also verify that $\mathrm{div}_\Gamma \partial_e \chi(0,\cdot) = 0$. This follows by calculating

$$\mathrm{div}_\Gamma \partial_e \chi(0,x) = \frac{1}{R}\left(\nabla_\Gamma \eta(x) \cdot (N \times x) + \eta(x) \mathrm{div}_\Gamma (N \times x)\right),$$

by the fact that η depends only on x_3, one sees that the first term is some scalar function multiplied by $P_\Gamma(x) N \cdot (N \times x)$, which vanishes. For the second term, one calculates, by extending to a small neighbourhood of the surface (as in the definition of surface derivatives),

$$\mathrm{div}_\Gamma (N \times x) = \sum_{i=1}^{3} \underline{D}_i (N \times x)_i = \sum_{i,j=1}^{3} \left(\delta_{ij} - \frac{x_i x_j}{R^2}\right) \partial_j (N \times x)_i \,.$$

We see that this vanishes, since $\delta_{ij} \partial_j (N \times x)_i = 0$ for any $i, j = 1, 2, 3$, and

$$\sum_{i=1}^{3} \frac{x_i x_j}{R^2} \partial_j (N \times x)_i = \sum_{i=1}^{3} \frac{x_j}{R^2} \partial_j \left(x_i (N \times x)_i\right) = 0$$

for any $j = 1, 2, 3$.

D.0.0.2 A general χ

Since the set $\bigcup_{i=1}^N \mathcal{C}_i(p)$ is a finite union of points, we know there is a strictly positive distance separating each pair of points. It follows that we may assume that the family of sets $\bigcup_{i=1}^N \mathcal{C}_i(p+tq)$ for $(t,q) \in [0,1] \times \mathcal{B}$ also satisfy this condition, and set $\epsilon > 0$ to be the smallest separation between the points of $\bigcup_{i=1}^N \mathcal{C}_i(p+tq)$ - that is

$$\epsilon = \inf_{(t,q) \in [0,1] \times \mathcal{B}} \inf_{x \in \bigcup_{i=1}^N \mathcal{C}_i(p+tq)} \inf_{y \in \bigcup_{i=1}^N \mathcal{C}_i(p+tq), y \neq x} |x - y|.$$

Definition D.0.1 (Equation (2.6) [71]). *We define the vector surface curl of a C^1 function $\psi \colon \Gamma \to \mathbb{R}$ by*

$$\mathrm{curl}_\Gamma \psi := \nu \times \nabla_\Gamma \psi.$$

Definition D.0.2. *Given $\delta \in (0, \epsilon)$, define $\mathcal{V} \colon [0,1] \times \mathcal{B} \times \Gamma \to \mathbb{R}^3$ by*

$$\mathcal{V} := \mathrm{curl}_\Gamma \psi$$

where for each $(t,q) \in [0,1] \times \mathcal{B}$, $x \in \bigcup_{i=1}^N \mathcal{C}_i(p+tq)$, the function $\psi \colon [0,1] \times \mathcal{B} \times \Gamma \to \mathbb{R}$ is given by

$$\psi(t,q,y) = \eta(|x-y|) y \cdot \left(\partial_s \left(\phi_i(p+sq,\cdot) \circ \phi_i(p+tq,\cdot)^{-1}(y)\right)|_{s=t} \times \nu(x)\right)$$

for $y \in \Gamma \cap B_{\epsilon/2}(x)$, otherwise $\psi = 0$, where $\eta \colon \mathbb{R} \to \mathbb{R}$ is a C^{k+1}-smooth cut off function such *that*

$$\begin{cases} \eta(s) = 1 & |s| \le \delta/4, \\ \eta(s) = 0 & |s| \ge \delta/2. \end{cases}$$

Example D.0.3. *We now give a calculation of* $\partial_s \left(\phi_i(p+sq,\cdot) \circ \phi_i(p+tq,\cdot)^{-1}(y)\right)|_{s=t}$. *For simplicity, we set* $p = 0$ *and* $t = 0$ *and neglect any* i *subscripts.*

Let $q = (\alpha, \tau) \in \mathbb{R} \times T_{X_\mathcal{G}}$. *We then have*

$$\phi(sq, x) = R_T(s\tau) R_n(s\alpha) x,$$

therefore

$$\partial_s \left(\phi(sq, x)\right)|_{s=0} = \left(\nu(X_\mathcal{G}) \times \tau\right) \times x + \alpha \left(\nu(X_\mathcal{G}) \times x\right).$$

It is clear that the first term corresponds to the translation and the second term the rotation. Without loss of generality we may keep $t = 0$ *(by changing* p *to be* $p + tq$*) then we calculate, writing* $p = (p_1, p_2) \in \mathbb{R} \times T_{X_\mathcal{G}}$,

$$\phi(p+sq,\cdot) \circ \phi(p,\cdot)^{-1}(x) = R_T(p_2 + s\tau) R_n(p_1 + s\alpha) R_n(p_1)^{-1} R_T(p_2)^{-1} x.$$

One then finds that

$$\begin{aligned} \partial_s \left(\phi(p+sq,\cdot) \circ \phi(p,\cdot)^{-1}(x)\right)|_{s=0} =& \partial_s \left(R_T(p_2 + s\tau)\right)|_{s=0} R_T(p_2)^{-1} x \\ &+ \alpha R_T(p_2) \left(\nu(X_\mathcal{G}) \times \left(R_n(p_1)^{-1} R_T(p_2)^{-1} x\right)\right), \end{aligned}$$

where

$$\begin{aligned} \partial_s \left(R_T(p_2 + s\tau)\right)|_{s=0} y =& -\frac{p_2 \cdot \tau}{|p_2|} \sin(|p_2|) y \\ &+ (1 - \cos(|p_2|)) \left(\left(\nu(X_\mathcal{G}) \times \left(\frac{\tau}{|p_2|} - \frac{\tau \cdot p_2}{|p_2|^3} p_2\right)\right) \cdot y\right) \tilde{\tau}(0) \\ &+ (1 - \cos(|p_2|)) \left(\tilde{\tau}(0) \cdot y\right) \left(\nu(X_\mathcal{G}) \times \left(\frac{\tau}{|p_2|} - \frac{\tau \cdot p_2}{|p_2|^3} p_2\right)\right) \\ &+ \frac{p_2 \cdot \tau}{|p_2|} \sin(|p_2|) \left(\tilde{\tau}(0) \cdot y\right) \tilde{\tau}(0) \\ &+ \sin(|p_2|) \left(\left(\nu(X_\mathcal{G}) \times \left(\frac{\tau}{|p_2|} - \frac{\tau \cdot p_2}{|p_2|^3} p_2\right)\right) \times y\right) \\ &+ \frac{p_2 \cdot \tau}{|p_2|} \cos(|p_2|) \left(\tilde{\tau}(0) \times y\right), \end{aligned}$$

where $\tilde{\tau}(0) := \nu(X_\mathcal{G}) \times \frac{p_2}{|p_2|}$. *Notice, if* $\tau = \beta p_2$, $\frac{\tau}{|p_2|} - \frac{\tau \cdot p_2}{|p_2|^3} p_2 = 0$ *and*

$$\partial_s \left(R_T(p_2 + s\tau)\right)|_{s=0} y = \beta |p_2| \sin(|p_2|) (\tilde{\tau}(0) \times y) \times \tilde{\tau}(0) - \beta |p_2| \cos(|p_2|) (\tilde{\tau}(0) \times y),$$

whereas, if $\tau \cdot p_2 = 0$,

$$\partial_s \left(R_T(p_2 + s\tau)\right)|_{s=0} y = (1 - \cos(|p_2|)) \left(\left(\nu(X_{\mathcal{G}}) \times \left(\frac{\tau}{|p_2|}\right)\right) \cdot y\right) \tilde{\tau}(0)$$
$$+ (1 - \cos(|p_2|)) \left(\tilde{\tau}(0) \cdot y\right) \left(\nu(X_{\mathcal{G}}) \times \left(\frac{\tau}{|p_2|}\right)\right)$$
$$+ \sin(|p_2|) \left(\left(\nu(X_{\mathcal{G}}) \times \left(\frac{\tau}{|p_2|}\right)\right) \times y\right).$$

Lemma D.0.4. *The function* $\mathcal{V}$ *given in Definition D.0.2 satisfies:*

- $\mathcal{V} \in C^k$,
- $div_\Gamma \mathcal{V} = 0$,
- $\mathcal{V}(t, 0, x) = 0$ *for all* $(t, x) \in [0, 1] \times \Gamma$,
- *for each* $i = 1, ..., N$, $\mathcal{V}(t, q, \cdot) = \partial_s \left(\phi_i(p + tq, \cdot) \circ \phi_i(p + sq, \cdot)^{-1}\right)|_{s=t}$ *on* $\mathcal{C}_i(p + tq)$, *for each* $(t, q) \in [0, 1] \times \mathcal{B}$,
- $\partial_e \mathcal{V}(t, 0, x) = \mathcal{V}(0, e, x)$ *for all* $t \in [0, 1]$, $e \in \prod_{i=1}^N \left(T_{X_{\mathcal{G}_i}} \times \mathbb{R}\right)$ *and* $x \in \Gamma$.

Proof. Smoothness and that $\mathcal{V}(\cdot, 0, \cdot)$ vanishes is clear by construction, divergence free follows from $\mathcal{V}$ being the curl of another function [71, Lemma 2.1]. For the point conditions we evaluate at $y \in \Gamma$ such that $|x - y| < \frac{\delta}{4}$ for some $x \in \mathcal{C}_i(p + tq)$,

$$\begin{aligned}
\text{curl}_\Gamma \psi(t, q, y) =& \text{curl}_\Gamma \left(y \cdot \left(\partial_s \left(\phi_i(p + sq, \cdot) \circ \phi_i(p + tq, \cdot)^{-1}(y)\right)|_{s=t}\right) \times \nu(x)\right) \\
=& \nu(y) \times \left(\nabla_\Gamma y \cdot \left(\partial_s \left(\phi_i(p + sq, \cdot) \circ \phi_i(p + tq, \cdot)^{-1}(y)\right)|_{s=t} \times \nu(x)\right)\right)
\end{aligned}$$

for each $(t, q) \in [0, 1] \times \mathcal{B}$, $i = 1, ..., N$. Which upon evaluation of at any $x \in \mathcal{C}_i(p + tq)$, $(t, q) \in [0, 1] \times \mathcal{B}$, $i = 1, ..., N$, leaves us with

$$\text{curl}_\Gamma \psi(t, q, x) = \partial_s \left(\phi_i(p + sq, \cdot) \circ \phi_i(p + tq, \cdot)^{-1}\right)|_{s=t}(x).$$

The final condition takes a little bit of work. We show the condition near the 'special points' of $\bigcup_{i=1}^N \mathcal{C}_i(p)$. Given $i = 1, ..., N$, for $x \in \mathcal{C}_i(p)$ and y near x, we see that

$$\begin{aligned}
\partial_e \mathcal{V}(t, 0, y) =& \partial_s \mathcal{V}(t, se, y)|_{s=0} \\
=& \partial_s \left(\mathcal{V}(t, se, \phi_i(p + se, \cdot) \circ \phi_i(p, \cdot)^{-1}(y))\right)|_{s=0} \\
&+ \partial_s \left(\mathcal{V}(t, se, x) - \mathcal{V}(t, se, \phi_i(p + se, \cdot) \circ \phi_i(p, \cdot)^{-1}(y))\right)|_{s=0} \\
=& \partial_s \left(\mathcal{V}(t, se, \phi_i(p + se, \cdot) \circ \phi_i(p, \cdot)^{-1}(y))\right)|_{s=0} \\
&+ \partial_s \left(\mathcal{V}(t, se, y) - \mathcal{V}(t, se, \phi_i(p + se, \cdot) \circ \phi_i(p, \cdot)^{-1}(y))\right)|_{s=0}.
\end{aligned}$$

This first term we may see is equal to $\mathcal{V}(0,e,x)$, for the remaining terms,

$$\begin{aligned}&\partial_s\left(\mathcal{V}(t,se,\phi_i(p+se,\cdot)\circ\phi_i(p,\cdot)^{-1}(y))\mathcal{V}(t,se,y)\right)|_{s=0}\\&\quad=\partial_s\left(\nabla_\Gamma\mathcal{V}(t,se,y)\cdot\left(\phi_i(p+se,\cdot)\circ\phi(p,\cdot)^{-1}(y)\right)-y\right)|_{s=0},\end{aligned}$$

which we see vanishes due to the fact that $\nabla_\Gamma\mathcal{V}(\cdot,se,\cdot)\to 0$ as $s\to 0$ on $[0,1]\times\Gamma$ and also $\phi_i(p+se,\cdot)\circ\phi(p,\cdot)^{-1}(y)-y\to 0$ as $s\to 0$. □

We will construct χ in the following way.

Definition D.0.5.

1. *Let $\eta\colon[0,1]\times\mathcal{B}\times\Gamma\to\Gamma$ be the solution of the family of ODEs*

$$\partial_t\eta(t,q,x)=\mathcal{V}(t,q,\eta(t,q,x)),\ \ \eta(0,q,x)=x$$

for all $(q,x)\in\mathcal{B}\times\Gamma$.

2. *Let $\chi\colon\mathcal{B}\times\Gamma\to\Gamma$ by $\chi(q,x)=\eta(1,q,x)$ for all $(q,x)\in\mathcal{B}\times\Gamma$.*

It is clear by standard ODE theory [52] that η exists and is smooth, furthermore, it is clear that $\eta(1,q,\cdot)$ is a diffeomorphism.

Proposition D.0.6. *For each $e\in\prod_{i=1}^N\left(T_{X_{\mathcal{G}_i}}\times\mathbb{R}\right)$, the following formula holds*

$$\partial_e\chi(0,\cdot)=\mathcal{V}(0,e,\cdot)\quad\text{on}\quad\Gamma.$$

Proof. This follows from the properties of $\mathcal{V}$ in Lemma D.0.4. The smoothness of χ follows from the smoothness of $\mathcal{V}$ and standard ODE theory [52], as does the existence and smoothness of an inverse. The condition that $\mathcal{V}(\cdot,0,\cdot)=0$ gives that $\chi(0,\cdot)$ is the identity.

The condition $v\circ\chi(q,\cdot)^{-1}\in U(p+q)\iff v\in U(p)$ has three parts:

- $v\circ\chi(q,\cdot)\in H^2(\Gamma)\iff v\in H^2(\Gamma)$,
- $\fint_\Gamma v=\fint_\Gamma v\circ\chi(q,\cdot)$ for all $v\in H^2(\Gamma)$,
- $T(p+q)\left(v\circ\chi^{-1}\right)=T(p)v$ for all $v\in H^2(\Gamma)$.

The first condition follows from two applications of Lemma A.0.2 with $X=\chi(q,\cdot)$ and $X=\chi(q,\cdot)^{-1}$ and the smoothness of these maps. The second condition follows from the fact that $div_\Gamma\mathcal{V}=0$. The final condition follows from the point conditions on $\mathcal{V}$. By considering the ODE that η solves, we see that χ satisfies for each $i=1,...,N$,

$$\chi(q,\cdot)=\phi_i(p+q,\cdot)\circ\phi_i(p,\cdot)^{-1}\text{ on }\mathcal{C}_i(p),$$

which gives, recalling the definition of T in Definition 2.1.15,

$$\begin{aligned}
T(p+q)_i v &= v \circ \phi_i(p+q,\cdot)|_{\mathcal{C}_i} \\
&= v \circ \phi_i(p+q,\cdot) \circ \phi_i(p,\cdot)^{-1} \circ \phi_i(p,\cdot)|_{\mathcal{C}_i} \\
&= T(p)_i(v \circ \chi(q,\cdot)).
\end{aligned}$$

□

We now wish to calculate $\partial_e \chi(0,\cdot)$ on Γ.

Proposition D.0.7. *For each $e \in \prod_{i=1}^N \left(T_{X_{\mathcal{G}_i}} \times \mathbb{R}\right)$, $\partial_e \chi(0,\cdot) = \mathcal{V}(0,e,\cdot)$ on Γ.*

Proof. It is clear that $\partial_e \chi(0,\cdot) = \partial_e \eta(1,0,\cdot)$. From the ODE η solves, one may see that $\eta_e(t,x) := \partial_e \eta(t,0,x)$ for $(t,x) \in [0,1] \times \Gamma$ satisfies

$$\partial_t \eta_e(t,x) = \partial_e \mathcal{V}(t,0,\eta(t,0,x)) + \nabla_\Gamma \mathcal{V}(t,0,\eta(t,0,x))\eta_e(t,x),$$

for all $(t,x) \in [0,1] \times \Gamma$. Recall that $\mathcal{V}(t,0,x) = 0$ for all $(t,x) \in [0,1] \times \Gamma$, so the second term in the above ODE vanishes and one has that $\eta(t,0,x) = x$ for all $(t,x) \in [0,1] \times \Gamma$. By applying the final condition of Lemma D.0.4, one has that

$$\partial_t \eta_e(t,x) = \mathcal{V}(0,e,x),$$

hence $\partial_e \chi(0,\cdot) = \eta_e(1,\cdot) = \mathcal{V}(0,e,\cdot)$ on Γ. □

Bibliography

[1] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, vol. 140 of Pure and Applied Mathematics (Amsterdam), Elsevier/Academic Press, Amsterdam, second ed., 2003.

[2] M. Alkämper, A. Dedner, R. Klöfkorn, and M. Nolte, *The DUNE-ALUGrid module.*, Archive of Numerical Software, 4 (2016), pp. 1–28.

[3] D. Arnold, R. Falk, and R. Winther, *Mixed finite element methods for linear elasticity with weakly imposed symmetry*, Math. Comp., 76 (2007), pp. 1699–1723.

[4] A. H. Bahrami and G. Hummer, *Formation and stability of lipid membrane nanotubes*, ACS Nano, 11 (2017), pp. 9558–9565. PMID: 28873296.

[5] J. W. Barrett, H. Garcke, and R. Nürnberg, *Finite element methods for fourth order axisymmetric geometric evolution equations*, J. Comput. Phys., 376 (2019), pp. 733 – 766.

[6] ——, *Stable approximations for axisymmetric Willmore flow for closed and open surfaces*, arXiv preprints arXiv:1911.01132, (2019).

[7] ——, *Variational discretization of axisymmetric curvature flows*, Numerische Mathematik, 141 (2019), pp. 791–837.

[8] A.-F. Bitbol, D. Constantin, and J.-B. Fournier, *Membrane-Mediated Interactions*, Springer International Publishing, Cham, 2018, pp. 311–350.

[9] M. Blatt, A. Burchardt, A. Dedner, C. Engwer, J. Fahlke, B. Flemisch, C. Gersbacher, C. Gräser, F. Gruber, C. Grüninger, et al., *The distributed and unified numerics environment, version 2.4*, Archive of Numerical Software, 4 (2016), pp. 13–29.

[10] D. Boffi, F. Brezzi, and M. Fortin, *Mixed finite element methods and applications*, vol. 44, Springer, 2013.

[11] J. H. Bramble, J. E. Pasciak, and O. Steinbach, *On the stability of the L^2 projection in $H^1(\Omega)$*, Math. Comp., 71 (2002), pp. 147–156.

[12] K. Brazda, L. Lussardi, and U. Stefanelli, *Existence of varifold minimizers for the multiphase Canham-Helfrich functional*, arXiv preprints arXiv:1912.02614, (2019).

[13] E. Bretin, S. Masnou, and E. Oudet, *Phase-field approximations of the Willmore functional and flow*, Numer. Math., 131 (2015), p. 115–171.

[14] G. Buttazzo and S. A. Nazarov, *An optimization problem for the biharmonic equation with Sobolev conditions*, J. Math. Sci., 176 (2011), p. 786.

[15] P. B. Canham, *The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell*, J. Theor. Biol., 26 (1970), pp. 61–81.

[16] G. Carey and R. Krishnan, *Penalty approximation of Stokes flow*, Comput. Methods Appl. Mech. Engrg., 35 (1982), pp. 169 – 206.

[17] E. Casas, L^2 *estimates for the finite element method for the Dirichlet problem with singular data*, Numer. Math., 47 (1985), pp. 627–632.

[18] L. Church, A. Djurdjevac, and C. M. Elliott, *A domain mapping approach for elliptic equations posed on random bulk and surface domains*, Numer. Math., https://doi.org/10.1007/s00211-020-01139-7 (2020).

[19] P. G. Ciarlet, *Linear and Nonlinear Functional Analysis with Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2013.

[20] P. Ciarlet Jr, J. Huang, and J. Zou, *Some observations on generalized saddle-point problems*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 224–236.

[21] D. Cuvelier, N. Chiaruttini, P. Bassereau, and P. Nassoy, *Pulling long tubes from firmly adhered vesicles*, Europhys. Lett., 71 (2005), pp. 1015–1021.

[22] K. Deckelnick, G. Dziuk, and C. M. Elliott, *Computation of geometric partial differential equations and mean curvature flow*, Acta Numer., 14 (2005), p. 139?232.

[23] K. Deckelnick, J. Katz, and F. Schieweck, *A* C^1*-finite element method for the Willmore flow of two-dimensional graphs*, Math. Comp., 84 (2015), pp. 2617–2643.

[24] I. Derényi, F. Jülicher, and J. Prost, *Formation and interaction of membrane tubes*, Phys. Rev. Let., 88 (2002), p. 238101.

[25] S. Dharmavaram and L. E. Perotti, *A lagrangian formulation for interacting particles on a deformable medium*, Comput. Method. Appl. M., 364 (2020), p. 112949.

[26] M. P. Do Carmo, *Differential geometry of curves and surfaces: revised and updated second edition*, Courier Dover Publications, 2016.

[27] P. G. Dommersnes and J.-B. Fournier, *The many-body problem for anisotropic membrane inclusions and the self-assembly of "saddle" defects into an "egg carton"*, Biophys. J., 83 (2002), pp. 2898 – 2905.

[28] G. Dziuk and C. M. Elliott, *Finite element methods for surface PDEs*, Acta Numer., 22 (2013), pp. 289–396.

[29] S. Eichmann, *The Helfrich boundary value problem*, Calc. Var. Partial Dif., 58 (2019).

[30] ——, *Lower-semicontinuity for the Helfrich problem*, arxiv preprints arXiv:1908.11738, (2019).

[31] C. M. Elliott, D. A. French, and F. A. Milner, *A second order splitting method for the Cahn-Hilliard equation*, Numerische Mathematik, 54 (1989), pp. 575–590.

[32] C. M. Elliott, H. Fritz, and G. Hobbs, *Small deformations of Helfrich energy minimising surfaces with applications to biomembranes*, Math. Models Methods Appl. Sci., 27 (2017), pp. 1547–1586.

[33] ——, *Second order splitting for a class of fourth order equations*, Math. Comp., 88 (2019), pp. 2605–2634.

[34] C. M. Elliott, C. Gräser, G. Hobbs, R. Kornhuber, and M.-W. Wolf, *A variational approach to particles in lipid membranes*, Arch. Ration. Mech. Anal., 222 (2016), pp. 1011–1075.

[35] C. M. Elliott and L. Hatcher, *Domain formation via phase separation for spherical biomembranes with small deformations*, European Journal of Applied Mathematics, (2020), p. 1–26. (FirstView article).

[36] C. M. Elliott, L. Hatcher, and P. J. Herbert, *Small deformations of spherical biomembranes*, in The Role of Metrics in the Theory of Partial Differential Equations, vol. 85 of Advanced Studies in Pure Mathematics, Tokyo, Japan, 2020, Mathematical Society of Japan, pp. 39–61. (to appear).

[37] C. M. Elliott and P. J. Herbert, *A formula for membrane mediated point particle interactions on near spherical biomembranes*, (2020). (submitted).

[38] ——, *Second order splitting of a class of fourth order pdes with point constraints*, Math. Comp., 89 (2020), pp. 2613–2648.

[39] C. M. Elliott and B. Stinner, *Modeling and computation of two phase geometric biomembranes using surface finite elements*, J. Comput. Phys., 229 (2010), pp. 6585 – 6612.

[40] ——, *A surface phase field model for two-phase biological membranes*, SIAM J. Appl. Math., 70 (2010), pp. 2904–2928.

[41] ——, *Computation of two-phase biomembranes with phase dependentmaterial parameters using surface finite elements*, Comm. Comput. Phys., 13 (2013), p. 325–360.

[42] A. Ern and J. Guermond, *Theory and Practice of Finite Elements*, vol. 159 of Applied Mathematical Sciences, Springer Verlag, New York, 2004.

[43] M. Fei and Y. Liu, *Phase-field approximation of the Willmore flow*, arXiv preprints arXiv:1904.11139, (2019).

[44] J.-B. Fournier and P. Galatola, *High-order power series expansion of the elastic interaction between conical membrane inclusions*, Eur. Phys. J. E, 38 (2015), p. 86.

[45] H. Garcke and R. Nürnberg, *Structure-preserving discretizations of gradient flows for axisymmetric two-phase biomembranes*, IMA J. Numer. Anal., (2020).

[46] H.-H. Gerdes and R. N. Carvalho, *Intercellular transfer mediated by tunneling nanotubes*, Curr. Opin. Cell Biol., 20 (2008), pp. 470 – 475.

[47] V. Girault and P. A. Raviart, *Finite element methods for Navier-Stokes equations*, vol. 5 of Springer Series in Computational Mathematics, Springer-Verlag, 1986.

[48] M. Goulian, R. Bruinsma, and P. Pincus, *Long-range forces in heterogeneous fluid membranes*, Europhys. Lett., 22 (1993), pp. 145–150.

[49] N. Gov, *Guided by curvature: Shaping cells by coupling curved membrane proteins and cytoskeletal forces*, Philos. T. R. Soc. B, 373 (2018).

[50] C. Gräser and T. Kies, *On differentiability of the membrane-mediated mechanical interaction energy of discrete-continuum membrane-particle models*, arXiv preprints arXiv:1711.11192, (2017).

[51] ——, *Discretization error estimates for penalty formulations of a linearized Canham–Helfrich-type energy*, IMA J. Numer. Anal., 39 (2019), pp. 626–649.

[52] P. Hartman, *Ordinary differential equations*, vol. 38 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.

[53] W. Helfrich, *Elastic properties of lipid bilayers: theory and possible experiments*, Z. Naturforsch. C, 28 (1973), pp. 693–703.

[54] W. M. Henne, E. Boucrot, M. Meinecke, E. Evergren, Y. Vallis, R. Mittal, and H. T. McMahon, *FCHo proteins are nucleators of clathrin-mediated endocytosis*, Science, 328 (2010), pp. 1281–1284.

[55] W. M. Henne, H. M. Kent, M. G. Ford, B. G. Hegde, O. Daumke, P. J. G. Butler, R. Mittal, R. Langen, P. R. Evans, and H. T. McMahon, *Structure and*

analysis of FCHo2 F-BAR domain: A dimerizing and membrane recruitment module that effects membrane curvature, Structure, 15 (2007), pp. 839 – 852.

[56] G. Hobbs, *Particles and biomembranes: a variational PDE approach*, PhD thesis, The University of Warwick Mathematics Institute, 2016.

[57] R. B. Kellogg and B. Liu, *A finite element method for the compressible Stokes equations*, SIAM J. Numer. Anal., 33 (1996), pp. 780–788.

[58] T. Kies, *Gradient methods for membrane-mediated particle interactions*, PhD thesis, Institut für Mathematik, Freie Universität Berlin, 2019.

[59] K. Kim, J. Neu, and G. Oster, *Curvature-mediated interactions between membrane proteins*, Biophy. J., 75 (1998), pp. 2274 – 2291.

[60] ——, *Effect of protein shape on multibody interactions between membrane inclusions*, Phys. Rev. E, 61 (2000), pp. 4281–4285.

[61] B. Kovács, B. Li, and C. Lubich, *A convergent evolving finite element algorithm for Willmore flow of closed surfaces*, arXiv preprints arXiv:2007.15257, (2020).

[62] E. Kuwert and R. Schätzle, *The Willmore flow with small initial energy*, J. Differ. Geom., 57 (2001), pp. 409–441.

[63] K. Larsson and M. G. Larson, *A continuous/discontinuous Galerkin method and a priori error estimates for the biharmonic problem on surfaces*, Math. Comp., 86 (2017), pp. 2613–2649.

[64] Z. Li, B. Anvari, M. Takashima, P. Brecht, J. H. Torres, and W. E. Brownell, *Membrane tether formation from outer hair cells with optical tweezers*, Biophys. J., 82 (2002), pp. 1386–1395.

[65] L. Marzo, K. Gousset, and C. Zurzolo, *Multifaceted roles of tunneling nanotubes in intercellular communication*, Front. Physiol., 3 (2012), p. 72.

[66] R. Mathias, S. Gabriel, and L. Tony, *Free Energy Computations: A Mathematical Perspective*, World Scientific, 2010.

[67] H. T. McMahon and J. L. Gallop, *Membrane curvature and mechanisms of dynamic cell membrane remodelling*, Nature, 438 (2005), p. 590.

[68] A. Mondino and C. Scharrer, *Existence and regularity of spheres minimising the Canham–Helfrich energy*, Arch. Ration. Mech. Anal., 236 (2020), p. 1455–1485.

[69] J. Necas, *Direct methods in the theory of elliptic equations*, Springer Science & Business Media, 2011.

[70] K. R. Porter, *Observations on a submicroscopic basophilic component of cytoplasm*, J. Exp. med., 97 (1953), pp. 727–750.

[71] A. Reusken, *Stream function formulation of surface Stokes equations*, IMA J. Numer. Anal., 40 (2020), pp. 109–139.

[72] A. Roux, *The physics of membrane tubes: soft templates for studying cellular membranes*, Soft Matter, 9 (2013), pp. 6726–6736.

[73] A. Rätz and M. Röger, *A new diffuse-interface approximation of the Willmore flow*, arXiv preprints arXiv:1910.14373, (2019).

[74] A. H. Schatz, *A weak discrete maximum principle and stability of the finite element method in l_∞ on plane polygonal domains. i*, Math. Comp., 34 (1980), pp. 77–91.

[75] R. Schätzle, *The Willmore boundary problem*, Calc. Var. Parital Dif., 37 (2010), pp. 275–302.

[76] Y. Schweitzer and M. M. Kozlov, *Membrane-mediated interaction between strongly anisotropic protein scaffolds*, PLOS Comput. Biol., 11 (2015), pp. 1–17.

[77] M. Sun, J. S. Graham, B. Hegedüs, F. Marga, Y. Zhang, G. Forgacs, and M. Grandbois, *Multiple membrane tethers probed by atomic force microscopy*, Biophys. J., 89 (2005), pp. 4320–4329.

[78] T. R. Weikl, *Membrane-mediated cooperativity of proteins*, Annu. Rev. Phys. Chem., 69 (2018), pp. 521–539.

[79] T. R. Weikl, M. M. Kozlov, and W. Helfrich, *Interaction of conical membrane inclusions: Effect of lateral tension*, Phys. Rev. E, 57 (1998), pp. 6988–6995.

[80] T. J. Willmore, *Note on embedded surfaces*, An. Sti. Univ.“Al. I. Cuza” Iasi Sect. I a Mat.(NS) B, 11 (1965), pp. 493–496.

[81] C. Yolcu, R. C. Haussman, and M. Deserno, *The effective field theory approach towards membrane-mediated interactions between particles*, Adv. Colloid Interfac., 208 (2014), pp. 89 – 109. Special issue in honour of Wolfgang Helfrich.

[82] A. Zhiliakov, Y. Wang, A. Quaini, M. Olshanskii, and S. Majd, *Experimental validation of a phase-field model to predict coarsening dynamics of lipid domains in multicomponent membranes*, arXiv preprints arXiv:2006.14125, (2020).